

A METAPOPOPULATION MODEL WITH MARKOVIAN LANDSCAPE DYNAMIC

R. McVINISH¹, P.K. POLLETT and Y.S. CHAN

School of Mathematics and Physics, University of Queensland

ABSTRACT. We study a variant of Hanski's incidence function model that allows habitat patch characteristics to vary over time following a Markov process. The widely studied case where patches are classified as either suitable or unsuitable is included as a special case. For large metapopulations, we determine a recursion for the probability that a given habitat patch is occupied. This recursion enables us to clarify the role of landscape dynamics in the survival of a metapopulation. In particular, we show that landscape dynamics affects the persistence and equilibrium level of the metapopulation primarily through its effect on the distribution of a local population's life span.

1. INTRODUCTION

A metapopulation is a collection of local populations of a single focal species occupying spatially distinct habitat patches. Much of the research on metapopulations has focussed on identifying and quantifying extinction risks, with mathematical modelling playing an important role. Levins [33] proposed the first model of a metapopulation, which, despite its many simplifying assumptions, provided a number of important insights [21]. The importance of spatial features such as landscape heterogeneity and patch connectivity to metapopulation persistence was demonstrated in subsequent research [22, 50, 45, 24] and by connections to interacting particle systems [34, 17, 18]. Of particular relevance to the current work is the Incidence Function Model (IFM) [22], which relates the colonisation and local extinction probabilities to landscape characteristics. This model has been used

PKP and RM are supported in part by the Australian Research Council (Discovery Grant DP150101459 and the ARC Centre of Excellence for Mathematical and Statistical Frontiers, CE140100049).

©2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

¹Corresponding author: email r.mcvinish@uq.edu.au

to study extinction risk and the effectiveness of conservation measures for a number of populations including the African lion (*Panthera leo*) in Kenya and Tanzania [16], the water vole (*Arvicola amphibius*) in the UK [36], and the prairie dog (*Cynomys ludovicianus*) in northern Colorado, USA [19].

While real landscapes are structured spatially, they also vary temporally. Landscape dynamics are known to play an important role in the persistence/extinction of a number of species [62]. As an example, Hanski [23] mentions the marsh fritillary butterfly (*Eurodryas aurinia*) whose host plant *Succisa pratensis* occurs in forest clearings that are between two and ten years old. The metapopulation of sharp-tailed grouse (*Tympanuchus phasianellus*), which occupies areas of grassland, is similarly affected by landscape dynamics [2, 20]. For this species, fire opens new grassland areas and prevents the encroachment of forests. Other examples include metapopulations of the perennial herb *Polygonella basiramia* [8] and metapopulations of the beetle *Stephanopachys linearis*, which breeds only in burned trees [53]. In these examples, the landscape dynamics are driven by secondary succession, and this is often the case regardless of whether the focal species is part of a seral community or the climax community.

Some authors [9, 63, 28] have attempted to deal with landscape dynamics by incorporating the time elapsed since the patch was created through the local extinction probability. A more widely used and studied approach incorporates landscape dynamics by allowing each patch to alternate between being suitable or unsuitable for supporting a local population. In its simplest form all patches are treated equally [31, 55, 65, 54]. A more general form used by DeWoody et al. [14] and Xu et al. [66] incorporates differences between patches in area and extinction rates. These studies demonstrate an important relationship between the time scale of metapopulation dynamics and landscape dynamics. When the habitat life span is too short, the metapopulation is unable to become established [31, 14]. Furthermore, ignoring landscape dynamics leads to inaccurate persistence criteria. In general, the persistence criteria is optimistic [55, 66], though not necessarily for species which are able to react to habitat destruction [54].

The main problem we see with the suitable/unsuitable classification is that it is too coarse. Treating patches as being unsuitable may be reasonable following a destructive event, but this approach is unable to handle typical environmental fluctuations in habitat size or quality. We note that earlier examples of modelling change in natural landscapes

with Markov chains used larger state spaces [4, Table 2]. A related but more subtle problem is that all two-state Markov chains, like those used to model landscape dynamics in the above papers, are reversible in the sense that the process appears the same when time is reversed [30, section 1.2]. The process of landscape succession following disturbance typically proceeds through a number of stages in a fixed order [51]. This is inconsistent with reversibility since for a reversible Markov chain any path which ultimately returns to the starting state must have the same probability regardless of whether this is traced in one direction or the other [30, section 1.5]. Furthermore, the transition kernel of a reversible Markov chain has a special structure which most Markov chains do not possess [30, Theorem 1.7].

In this paper we adopt a similar approach to the one proposed by Hanski [23] and incorporate landscape dynamics into the incidence function model by treating landscape characteristics as a stochastic process. To the authors' knowledge, this approach has not been subject to a detailed mathematical analysis. Specifically, we model the landscape characteristics as a Markov chain on a general state space. The Markov chain model facilitates the analysis while allowing a general state space avoids the issues raised with the suitable/unsuitable classification. Our primary aim is to understand the affect of landscape dynamics on the survival of the metapopulation. Previous studies using the suitable/unsuitable classification have shown that landscape dynamics affect the survival of large metapopulations only through the average life span of the local populations [14, 66]. It is natural to consider whether this behaviour holds for more general landscape dynamics. In order to address this question, we study the limiting behaviour of the metapopulation when the number of patches is large. Using an analysis similar to our earlier work [38, 39, 41], we show large metapopulations display a deterministic limit and asymptotic independence of local populations. All proofs are given in the Appendix.

2. MODEL DESCRIPTION

Hanski's IFM [22] describes the evolution of a metapopulation as a discrete-time Markov chain. In this model, the dynamics of the metapopulation are determined by characteristics of the habitat patches. These characteristics are the patch's location, a weight related to the size of the patch, and the probability that a local population occupying this patch survives a given period of time. The colonisation and extinction processes

which govern the metapopulation are specified by these characteristics. The local population of a patch will go extinct at the next time step with probability given by one minus the local survival probability. A patch that is currently empty will be colonised by individuals migrating from the occupied patches at the next time step with a probability that depends on its relative location to the occupied patches and the weight associated with those occupied patches. For a metapopulation comprising n patches, the presence or absence of the focal species at each of the patches determines the metapopulation's state.

Before describing the metapopulation dynamics more formally, we need to introduce some notation. The state of the metapopulation at time t is denoted by the binary vector $X_t^n = (X_{1,t}^n, \dots, X_{n,t}^n)$, where $X_{i,t}^n = 1$ if patch i is occupied at time t and $X_{i,t}^n = 0$ otherwise. The patch location, weight and survival probability of patch i are denoted by z_i , A_i and s_i , respectively. We let z^n , A^n and s^n denote the respective vectors of all n patch locations, weights and survival probabilities. Since we view the landscape as the result of some random process, we treat z^n , A^n and s^n as random vectors. Conditional on the state of the metapopulation at time t and the patch characteristics, the $X_{i,t+1}^n$ ($i = 1, \dots, n$) are independent with transition probabilities

$$\mathbb{P}(X_{i,t+1}^n = 1 \mid X_t^n, z^n, A^n, s^n) = s_i X_{i,t}^n + f\left(\sum_{j=1}^n X_{j,t}^n D(z_i, z_j) A_j\right) (1 - X_{i,t}^n), \quad (2.1)$$

where $D(z_i, z_j) = \exp(-\alpha \|z_i - z_j\|)$ for some $\alpha > 0$ and the function $f : [0, \infty) \rightarrow [0, 1]$ is the colonisation function. Hanski [22] takes the colonisation function to be $f(x) = \beta x^2 / (\gamma + \beta x^2)$ for some $\beta, \gamma > 0$ and assumes the survival probability s_i is a function of A_i . The sum $\sum_{j=1}^n X_{j,t}^n D(z_i, z_j) A_j$ is called the connectivity measure. It specifies how each occupied patch contributes to the potential colonisation of any given unoccupied patch i . Larger patches make a greater contribution to the connectivity measure as they are expected to have a greater capacity to produce propagules. These propagules then disperse to other patches based on their proximity so nearby patches contribute more to the connectivity measure.

Before incorporating landscape dynamics into this model, it will be useful to briefly discuss our approach to the analysis. The central idea is that as the number of habitat patches increases, the connectivity measure at each location converges to a deterministic limit in probability. This is based on scaling the connectivity measure with the number

of patches in the metapopulation and can be achieved in a number of ways. One way is if the total habitable area for the focal species is fixed. As the number of patches increases, the landscape becomes more fragmented. If all patches are of a comparable size, then A_i should be of the order n^{-1} and the connectivity measure for patch i can be expressed as

$$n^{-1} \sum_{j=1}^n X_{j,t}^n D(z_i, z_j) a_j, \quad (2.2)$$

where $A_j = n^{-1}a_j$. This type of scaling is discussed in Barbour et al. [6, section 5] in connection with approximating the stochastic model (2.1) by a deterministic difference equation. Alternatively, we might consider that each patch has a finite number of propagules that can be dispersed in a given period. These propagules are divided between the $n - 1$ potential destination patches based on their proximity. As $n \rightarrow \infty$, the propagules are divided between more patches. This introduces a factor of n^{-1} in the connectivity measure resulting in (2.2). Finally, in a similar spirit to [49], we could consider an increasing region with constant density of patch locations and scale the dispersal kernel so that $D_n(z_i, z_j) = \alpha_n^d \exp(-\alpha_n \|z_i - z_j\|)$, with $\alpha_n \rightarrow 0$. Provided the rates at which the region increases and α_n ‘balance’, the same limiting model will be obtained. Other scalings that result in a ‘weak law of large numbers’ for the connectivity measure, that is convergence in probability of the connectivity measure to a deterministic quantity, will lead to a similar limiting process.

We incorporate landscape dynamics into the incidence function model assuming that only the survival probabilities s_i and the patch weights a_i evolve over time and the patch locations remain static. The survival probability and weight of a patch are modelled using a single variable θ , which we call the *characteristic* of the patch, taking values in some set Θ . If the characteristic of patch i at time t is $\theta_{i,t}$, then the survival probability and weight are given by $s(\theta_{i,t})$ and $a(\theta_{i,t})$ respectively, where $s : \Theta \rightarrow [0, 1]$ and $a : \Theta \rightarrow [0, \infty)$. In this way, we allow dependence between the local survival probability and patch weight. If $a(\cdot)$ is an invertible function, then we can express the local survival probabilities as a function of patch weights as in Hanski [22]. Conditional on X_t^n , $\theta_t^n := (\theta_{1,t}, \dots, \theta_{n,t})$ and z^n , the $X_{i,t+1}^n$ ($i = 1, \dots, n$) are independent with transition probabilities

$$\mathbb{P}(X_{i,t+1}^n = 1 \mid X_t^n, \theta_t^n, z^n) = s(\theta_{i,t}) X_{i,t}^n + f \left(n^{-1} \sum_{j=1}^n X_{j,t}^n D(z_i, z_j) a(\theta_{j,t}) \right) (1 - X_{i,t}^n). \quad (2.3)$$

Equation (2.3) is a natural extension of Hanski's incidence function model allowing for a dynamic landscape. However, the form of the colonisation probability means that the characteristic of patch i does not affect the probability that it is colonised, and this may be undesirable in some applications. Moilanen and Nieminen [45] consider a connectivity measure where the size of the target patch increases the probability of colonisation. Here we allow the colonisation function to be a function of both the connectivity and the characteristic of patch to be colonised, that is,

$$\mathbb{P}(X_{i,t+1}^n = 1 \mid X_t^n, \theta_t^n, z^n) = s(\theta_{i,t})X_{i,t}^n + f\left(n^{-1} \sum_{j=1}^n X_{j,t}^n D(z_i, z_j) a(\theta_{j,t}); \theta_{i,t}\right) (1 - X_{i,t}^n). \quad (2.4)$$

A model with the connectivity measure proposed by Moilanen and Nieminen [45] is obtained by setting $f(x, \theta) = \bar{f}(b(\theta)x)$ for some functions $\bar{f} : [0, \infty) \rightarrow [0, 1]$ and $b : \Theta \rightarrow [0, \infty)$. Perhaps more importantly, model (2.4) allows for two effects that are not possible with model (2.3); phase structure and pulsed dispersal.

Suppose the colonisation and extinction phases alternate, with observations of the metapopulation made after the extinction phase. In this case the colonisation probability has the form $f(x, \theta) = s(\theta)\bar{f}(x, \theta)$ where $\bar{f} : [0, \infty) \times \Theta \rightarrow [0, 1]$. This type of phase structure has previously been used in [1, 13, 25, 39]. We note that if observations were instead taken after the colonisation phase, then the model would display the rescue effect [22].

Pulsed dispersal occurs when migration from a colonised patch is the result of the species response to a decline in habitat quality. A continuous time metapopulation model with suitable/unsuitable landscape dynamics is studied in [54]. Suppose for simplicity that $\Theta = \{1, 2\}$, where the state 1 indicates a suitable habitat patch and state 2 indicates an unsuitable habitat patch. Pulsed dispersal is achieved by taking $a(2) > a(1)$, that is an occupied patch that has recently become unsuitable will contribute more to the colonisation of other patches than an occupied patch that is still suitable. Taking $s(2) = f(\cdot, 2) = 0$ ensures the local population at the unsuitable patch becomes extinct with probability one and a patch cannot be recolonised while it is unsuitable.

3. ASYMPTOTIC BEHAVIOUR OF A SINGLE PATCH

The typical approach to incorporating landscape dynamics into a metapopulation model is to allow each patch to alternate between being suitable or unsuitable for supporting a local population according to some Markov chain, independently of the other patches and of the state of the metapopulation [31, 55, 65, 66, 54]. A natural extension is to model the temporal evolution of each patch characteristic by a Markov chain on a finite state space $\{1, \dots, m\}$ with a common transition probability matrix. However, the classification of habitat into one of a finite number of classes may be unnatural in some settings since, for example, it restricts patch areas to taking only finitely many values. To avoid this, we model the habitat dynamics as a Markov chain on a general state space with a common transition probability kernel.

We briefly recall some properties of Markov chains. Suppose Y_t is a Markov chain on the state space S and let Σ be the set of all events of interest (that is, a σ -field of subsets of S). The transition probability kernel P gives the probability that the chain moves from a point y to the set $A \in \Sigma$ in one time step:

$$P(y, A) = \mathbb{P}(Y_{t+1} \in A \mid Y_t = y). \quad (3.5)$$

When the Θ is finite, Σ is just the set of all subsets of S and the right hand side of (3.5) is just $\sum_{j \in A} P_{yj}$ where (P_{ij}) is the transition probability matrix of the Markov chain. Under certain weak conditions, the transition kernel has an invariant distribution π , that is

$$\pi(A) = \int_{\Theta} P(x, A) \pi(dx), \quad \text{for all } A \in \Sigma.$$

When the S is finite, so P is a transition probability matrix, the invariant distribution is simply a distribution π on S such that $\pi P = \pi$ and we write $\pi(A) = \sum_{i \in A} \pi_i$. In any case, if π is the distribution of Y_1 , then Y_t will also have this distribution for all t and the Markov chain is said to be stationary or in equilibrium.

Stationary Markov chains with a state space of only two elements have the rather special property of reversibility. Informally, this means the process will look the same if the direction of time is reversed. More precisely, a stationary Markov chain is reversible with respect to π if

$$\int_A \pi(dx) P(x, B) = \int_B \pi(dx) P(x, A), \quad \text{for all } A, B \in \Sigma.$$

For a finite state space, the condition for reversibility is simply $\pi_i P_{ij} = \pi_j P_{ji}$, for all $i, j \in S$.

Our analysis makes use of the dual transition kernel, a concept related to reversibility. If P has invariant distribution π , then there is a transition kernel P^* called a *dual of P with respect to π* satisfying

$$\int_A \pi(dx) P(x, B) = \int_B \pi(dx) P^*(x, A), \quad \text{for all } A, B \in \Sigma \quad (3.6)$$

(Theorem 7.1 of Appendix B). For a finite state space, the dual transition probability matrix is the probability transition matrix satisfying $\pi_i P_{ij} = \pi_j P_{ji}^*$, for all $i, j \in S$. This expression provides a means of constructing P^* from P and π .

As previously noted, the landscape might be viewed as the result of some random process. Here we assume that the $(\theta_{i,0}, z_i)$, $i = 1, \dots, n$, are independent and identically distributed. The marginal distribution of the patch locations is assumed to be supported on $\Omega \subset \mathbb{R}^d$ and have a probability density function which we denote by ζ . Under mild conditions on the transition kernel of the patch characteristic, landscapes that have existed for a long time should at least be approximately stationary in the sense that the distribution of $\theta_{i,t}$ should converge to its invariant distribution as $t \rightarrow \infty$. For Markov chains on general state spaces, convergence to the invariant distribution is ensured by the technical condition of positive Harris recurrence. When the state space is finite, such as for the suitable/unsuitable classification, positive Harris recurrence holds if the Markov chain is irreducible, that is each state can be reached from any other, and aperiodic. The following result shows this implies independence of the patch location and characteristic.

Lemma 3.1. *Suppose that Markov chain $(\theta_{i,t}, t \geq 0)$ taking values in Θ is positive Harris and aperiodic with invariant distribution π . For any i , the distribution of $(\theta_{i,t}, z_i)$ converges as t goes to infinity. Furthermore, $\theta_{i,t}$ and z_i are asymptotically independent in the sense that*

$$\lim_{t \rightarrow \infty} \mathbb{P}(\theta_{i,t} \in A, z_i \in B) = \pi(A) \int_B \zeta(z)(dz), \quad (3.7)$$

for any measurable $A \subset \Theta$ and $B \subset \Omega$.

The main result of this section concerns the behaviour of local populations. This result depends on being able to establish convergence in probability of the connectivity measure at each location as the number of habitat patches increases. A sequence of

random variables $\{Y^n\}_{n=1}^\infty$ is said to converge in probability to a random variable Y if, for any $\epsilon > 0$, $\mathbb{P}(\|Y^n - Y\| \geq \epsilon) \rightarrow 0$. Convergence in probability is denoted $Y^n \xrightarrow{p} Y$. We are now able to state our results on the behaviour of large metapopulations with Markovian landscape dynamics.

Theorem 3.2. *Suppose that*

$$\mathbb{P}(X_{i,0}^n = 1 \mid \theta_{i,0} = \theta, z_i = z) = q_0(\theta, z), \quad (3.8)$$

for some function $q_0 : \Theta \times \Omega \rightarrow [0, 1]$. Under the assumptions given in Appendix A, if $X_{i,0}^n \xrightarrow{p} X_{i,0}$, then $X_{i,t}^n \xrightarrow{p} X_{i,t}$ for all $t \geq 0$, where the transition probability for $X_{i,t}$ is

$$\mathbb{P}(X_{i,t+1} = 1 \mid X_{i,t}, \theta_{i,t} = \theta, z_i) = s(\theta)X_{i,t} + f(\psi_t(z_i); \theta)(1 - X_{i,t}) \quad (3.9)$$

and

$$\psi_t(z) = \int_{\Omega} D(z, \tilde{z}) \left(\int_{\Theta} a(\theta) q_t(\theta, \tilde{z}) \pi(d\theta) \right) \zeta(\tilde{z}) d\tilde{z}, \quad (3.10)$$

$$q_{t+1}(\theta, z) = \int_{\Theta} s(\eta) q_t(\eta, z) P^*(\theta, d\eta) + \int_{\Theta} f(\psi_t(z); \eta) (1 - q_t(\eta, z)) P^*(\theta, d\eta). \quad (3.11)$$

When Θ is a finite state space, equations (3.10) and (3.11) can be expressed as

$$\psi_t(z) = \int_{\Omega} D(z, \tilde{z}) \left(\sum_j a(j) q_t(j, \tilde{z}) \pi_j \right) \zeta(\tilde{z}) d\tilde{z}, \quad (3.12)$$

$$q_{t+1}(i, z) = \sum_j s(j) q_t(j, z) P_{ij}^* + \sum_j f(\psi_t(z); j) (1 - q_t(j, z)) P_{ij}^*. \quad (3.13)$$

The conditions for Theorem 3.2 to hold are given in a very general form in Appendix A. In the context of a finite state space for the landscape characteristic, sufficient conditions for Theorem 3.2 to hold are as follows: (i) the landscape dynamics are irreducible, aperiodic and stationary, (ii) the support for the patch locations is bounded, and (iii) the colonisation function is Lipschitz continuous for each state of the landscape characteristic. These assumptions are very mild. As previously mentioned, landscapes that have existed for a long time should at least be approximately stationary. To the authors' knowledge, all colonisation functions used in practice are Lipschitz continuous.

This theorem states that the process $((X_{i,t}^n, \theta_{i,t}, z_i), t \geq 0)$ converges to a Markov chain and is asymptotically (as $n \rightarrow \infty$) independent of the rest of the metapopulation. It is possible to extend this result to the case of a finite collection of habitat patches so that the local population at each patch is independent of the local populations of all others in

this collection. Results of this type are sometimes referred to as ‘propagation of chaos’ [for example 32, Proposition 4.3].

The following two results provide the recursion (3.10) - (3.11) with a natural interpretation.

Theorem 3.3. *Let $((X_{i,t}, \theta_{i,t}, z_i), t \geq 0)$ be the Markov chain defined by (3.9) - (3.11). Then*

$$\mathbb{P}(X_{i,t} = 1 \mid \theta_{i,t} = \theta, z_i = z) = q_t(\theta, z)$$

for all $t \geq 0$.

Theorem 3.3 shows that $q_t(\theta, z)$ is the probability of a patch located at z with characteristic θ being occupied at time t in the large metapopulation.

Theorem 3.4. *Let h be a continuous function on $\Theta \times \Omega$. Under the assumptions of Theorem 3.2,*

$$\lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n X_{i,t} h(\theta_{i,t}, z_i) \xrightarrow{p} \int_{\Omega} \left(\int_{\Theta} h(\theta, z) q_t(\theta, z) \pi(d\theta) \right) \zeta(z) dz. \quad (3.14)$$

For a finite state space, the right hand side of (3.14) is $\int_{\Omega} \left(\sum_j h(j, z) q_t(j, z) \pi_j \right) \zeta(z) dz$. If $a(\theta)$ is a continuous function, then we can apply Theorem 3.4 to expression (2.2) to see that the connectivity measure of patch i at time t is approximated by $\psi_t(z_i)$ for n sufficiently large. Theorem 3.4 can also be used to show that $\int_{\Omega} \int_{\Theta} q_t(\theta, z) \pi(d\theta) \zeta(z) dz$ is a good approximation to the proportion of colonised patches in the metapopulation at time t when n is large.

4. EQUILIBRIUM OF A METAPOPULATION WITH PHASE STRUCTURE

When the number of patches is finite, the metapopulation described by (2.4) goes extinct in finite time with probability one. However, we have seen in the previous section that when the number of patches in the metapopulation is large, its temporal trajectory is closely tracked by the recursion (3.10)-(3.11). Therefore, if q_t is bounded away from zero for all $t \geq 0$, then the metapopulation may persist for a long time. On the other hand if q_t goes to zero, then the proportion of occupied patches converges to zero and the metapopulation will go extinct quickly. In this section, we provide a criterion to distinguish between these two scenarios.

Persistence criteria have been established for a number of population models. Chesson [11] and Metz and Gyllenberg [42] derive persistence criteria for structured metapopulation models, that is where the size of the local population at each patch is modelled, not just its presence or absence. Ovaskainen and Hanski [50] established persistence criteria for a deterministic version of the incidence function model and for the spatially realistic Levins' model. In their criteria, the quantity determining persistence factorises into two parts, one which is dependent on the species' dispersal kernel and the landscape, and the other which is a function of non-spatial species specific parameters. The criterion for the spatially realistic Levins' model has been extended to allow suitable/unsuitable patch dynamics [14, 66]. These criteria are closely related to the basic reproduction number in disease modelling [61, 15].

While our previous results hold under rather weak assumptions, to analyse the recursion (3.10)-(3.11) we need to impose some more restrictive assumptions. Two main assumptions are used in the analysis. The first is that the metapopulation has a phase structure. This assumption implies that the colonisation function has form

$$f(x; \theta) = s(\theta) \bar{f}(x; \theta) \quad (4.15)$$

for some function $\bar{f}(x; \theta) : [0, \infty) \times \Theta \rightarrow [0, 1]$. The second assumption is that for each $\theta \in \Theta$, $\bar{f}(\cdot, \theta)$ is concave. The assumption of a concave colonisation function essentially excludes the possibility of an Allee-like effect in the metapopulation [12]. Although it precludes the colonisation function used by Hanski [22], it is sufficiently weak to accommodate a wide range of functions including the one used in [25].

The persistence criterion for model (2.4) is expressed in terms of two quantities. The first quantity r_S is based on how the focal species reacts to fluctuations in the landscape characteristic and is defined by

$$r_S := \sum_{m=1}^{\infty} \mathbb{E} \left\{ \bar{f}'(0; \theta_0) \left[\prod_{n=0}^{m-1} s(\theta_n) \right] a(\theta_m) \right\},$$

where $(\theta_t, t \geq 0)$ is the Markov chain with transition kernel P and stationary distribution π . The second quantity depends on how the focal species disperses in the landscape and on the distribution of patch locations. It is given by the spectral radius $r(\mathcal{M})$ of the

bounded linear operator $\mathcal{M} : C(\Omega) \rightarrow C(\Omega)$ defined by

$$\mathcal{M}\phi(z) := \int_{\Omega} D(z, \tilde{z}) \phi(\tilde{z}) \zeta(\tilde{z}) d\tilde{z}, \quad \phi \in C(\Omega).$$

Note that this quantity is independent of the landscape dynamics.

Theorem 4.1. *Suppose the assumptions listed in Appendices A and C hold. If $r_S \times r(\mathcal{M}) \leq 1$, then recursion (3.12)-(3.13) has only the trivial fixed point $q(\theta, z) = 0$ for all $(\theta, z) \in \Theta \times \Omega$, and this fixed point is globally stable. If $r_S \times r(\mathcal{M}) > 1$, then recursion (3.10)-(3.11) has a unique non-zero fixed point. Furthermore, if $q_0(\theta, z) > 0$ for all $(\theta, z) \in \Theta \times \Omega$, then q_t converges to this non-zero fixed point.*

The assumptions from Appendix A were explained following Theorem 3.2 in the context of a finite state space for the landscape dynamics. The two main assumptions from Appendix C have already been mentioned at the beginning of this section, namely that the colonisation function has the form (4.15) and for each $\theta \in \Theta$, $\bar{f}(\cdot, \theta)$ is concave. When Θ is finite, the two other assumptions simplify. One is that $s(i) < 1$ for each $i \in \Theta$. For the final assumption define $\Theta_1 := \{j \in \Theta : s(j)f(x, j) = 0 \text{ for all } x \in [0, \infty)\}$. The set Θ_1 comprises those states of the landscape characteristic for which a patch cannot be colonised. The final assumption is that there is an $i \in \Theta \setminus \Theta_1$ and $j \in \Theta$ such that $a(j) > 0$ and $P_{ij} > 0$. This means that, given the landscape characteristic was in a state that allowed the patch to be colonised, there is positive probability that at the next time step the landscape characteristic will be in a state which makes a positive contribution to the colonisation of the other patches.

Theorem 4.1 is an example of the kind of dichotomy observed in other metapopulation models not displaying an Allee-like effect. Without the assumption that \bar{f} is concave, the condition $r_S \times r(\mathcal{M}) > 1$ still implies the existence of a non-zero fixed point, but there may be several non-zero fixed points in this case. Similarly, if \bar{f} is only locally concave in a neighbourhood of zero and $r_S \times r(\mathcal{M}) < 1$, then the extinction fixed point is still locally stable. That is, iterations of the recursion (3.12)-(3.13) will converge to the fixed point $q(\theta, z) = 0$ for all $(\theta, z) \in \Theta \times \Omega$ for sufficiently small initial conditions. This follows from the monotonicity property of the recursion (see the proof of Theorem 8.8). However, a non-zero fixed point may also exist in this case.

Despite the significant differences between the spatially realistic Levins' model with dynamic landscape studied in [14, 66] and the model studied here, there are still some important similarities in their persistence criteria. Firstly, the quantity determining persistence factorises into the product of two terms; one determined by the dispersal kernel of the species and the location of habitat patches, and the other by the species reaction to landscape dynamics and other non-spatial factors. If we suppose that all patches have constant area, so $a(\theta) = \bar{a}$ for all $\theta \in \Theta$, then the factor depending on the species' dispersal kernel is the leading eigenvalue of the matrix M with elements $M_{ij} := D(z_i, z_j)$. When the patch locations are a sample of independent random variables from the probability density function ζ , the leading eigenvalue of M can be shown, under certain conditions, to converge to the spectral radius of \mathcal{M} as $n \rightarrow \infty$. On restricting the model so that $\bar{f}(x; \theta) = \bar{f}(x)$, further similarities appear. We can now write

$$r_S = \bar{a} \bar{f}'(0) \sum_{m=0}^{\infty} \mathbb{E} \left(\prod_{n=0}^{m-1} s(\theta_n) \right).$$

Noting that $\mathbb{E}(\prod_{n=0}^{m-1} s(\theta_n))$ is the probability that a local population survives at least m extinction phases, we see that the quantity $\sum_{m=0}^{\infty} \mathbb{E}(\prod_{n=0}^{m-1} s(\theta_n))$ is the expected life span of a local population. For the spatially realistic Levins' model with dynamic landscape, the quantity determining persistence has a factor $(e + \beta)^{-1}$, where e is the extinction rate and β is the rate of patch destruction. Therefore, the factor $(e + \beta)^{-1}$ gives the expected life span of the local population. These similarities suggest a certain amount of robustness of the conclusions to the particular modelling choices made.

Restricting the model to $\bar{f}(x; \theta) = \bar{f}(x)$ also enables us to have a clearer understanding of how the landscape dynamics affect the equilibrium level of the metapopulation.

Theorem 4.2. *Suppose $\bar{f}(x; \theta) = \bar{f}(x)$. Let q^* be a fixed point of recursion (3.10)-(3.11). The quantity $\int_{\Theta} a(\theta) q^*(\theta, z) \pi(d\theta)$ depends on (π, P) only through the sequence*

$$\mathbb{E} \left(a(\theta_{m+1}) \prod_{n=1}^m s(\theta_n) \right), \quad m \geq 1. \quad (4.16)$$

Theorem 4.2 states that, when the metapopulation is in equilibrium, the expected area taken up by colonised patches, $\int_{\Theta} a(\theta) q^*(\theta, z) \pi(d\theta)$, depends on the landscape dynamics only through expected future contributions to the connectivity measure of a colonised patch during the local population's life span. This future contribution to the connectivity

measure may be interpreted as the number of propagules produced. If all patches have constant area, then the sequence (4.16) is a multiple of the tail probabilities of the life span of the local population. In that case, the probability that a patch at location z is colonised at equilibrium depends on the landscape dynamics only through the distribution of the local population's life span.

Theorem 4.3. *Suppose $\bar{f}(x; \theta) = \bar{f}(x)$. Let $(\theta_t, t \geq 0)$ and $(\tilde{\theta}_t, t \geq 0)$ denoted the Markov chains generated by (π, P) and $(\tilde{\pi}, \tilde{P})$, respectively, and let q^* and \tilde{q}^* be the fixed points of the respective recursions (3.12)-(3.13). If*

$$\mathbb{E} \left(a(\tilde{\theta}_{m+1}) \prod_{n=1}^m s(\tilde{\theta}_n) \right) \leq \mathbb{E} \left(a(\theta_{m+1}) \prod_{n=1}^m s(\theta_n) \right), \quad (4.17)$$

for all $m \geq 1$, then

$$\int_{\Theta} a(\tilde{\theta}) \tilde{q}^*(\theta, z) \tilde{\pi}(d\theta) \leq \int_{\Theta} a(\theta) q^*(\theta, z) \pi(d\theta), \quad (4.18)$$

for all $z \in \Omega$.

Theorem 4.3 seems intuitively obvious; landscape dynamics which yield greater expected contributions to the connectivity measure result in greater expected area taken up by the colonised patches in equilibrium. What is important is that inequality (4.17) must hold for the entire sequence so the expected contribution must be greater at all future times during the local population's life span. If all patches have constant area, then inequality (4.17) reduces to stating that the life span of the local population is larger, in the usual stochastic ordering [57, Section 1.A], under (π, P) than under $(\tilde{\pi}, \tilde{P})$.

Our final result identifies the Markov chain which makes the life span of the local population the longest for a given stationary distribution π .

Corollary 4.4. *If the Markov chain $(\theta_t, t \geq 0)$ is stationary, then*

$$\mathbb{E} \left(\prod_{n=0}^m s(\theta_n) \right) \leq \mathbb{E}(s(\theta_0)^{m+1}). \quad (4.19)$$

This implies that among Markov chains with the same stationary distribution π , the one which makes the life span of the local population the greatest is the one for which $\theta_0 = \theta_t$ for all $t \geq 0$, that is the static landscape. This result with Theorem 4.3 implies that a static landscape maximises the probability of a patch being occupied when the

patch areas are constant. We have not been able to show a similar result for the sequence (4.16). The landscape dynamics which maximises the sequence (4.16) would seem to be affected by how patch areas and local survival probabilities depend on the landscape dynamics. This is to be expected since the model incorporates the possibility of pulsed dispersal.

To conclude this section, we perform some simulations comparing the proportion of time a patch in the metapopulation is occupied with the limiting probability of patch occupancy determined by $\int_{\Theta} q^*(\theta, z) \pi(d\theta)$ where $q^*(\theta, z)$ is the fixed point of the recursion (3.10) - (3.11). All simulations are performed with constant patch areas $a(\theta) = 10$ for all $\theta \in \Theta$, $\bar{f}(x) = 1 - \exp(-x)$ and $D(z, \tilde{z}) = \exp(-\|z - \tilde{z}\|)$. To facilitate the presentation, we assume a one dimensional landscape. The patch locations are sampled from the uniform distribution on $[0, 10]$.

The survival probabilities $s_t = s(\theta_t)$ are modelled by the Markov chain studied in McKinlay and Borovkov [37]. This Markov chain is defined by

$$s_{t+1} = \begin{cases} s_t(1 - L_{t+1}), & \text{with probability } p(s_t), \\ s_t + (1 - s_t)R_{t+1}, & \text{with probability } 1 - p(s_t), \end{cases} \quad (4.20)$$

where $p : [0, 1] \rightarrow [0, 1]$, and $\{L_t\}$ and $\{R_t\}$ are sequences of independent and identically distributed random variables on $[0, 1]$ with distributions F_L and F_R , respectively. Two sample paths are plotted in Figure 1 for two choices of F_L and F_R with $p(x) = 10(x - 0.9)\mathbb{I}(x > 0.9)$. Although we do not prove that this process is not reversible, the plotted sample paths strongly suggest that it is not. Specifically, the process would not look the same if time were reversed since the large downward jumps in the trajectory would appear as large upwards jumps if time were reversed.

This Markov chain can provide a reasonable model of changes in habitat quality due to disturbance followed by a slow restoration as follows. Immediately after a disturbance, the habitat is low quality so the local survival probability is small. As time progresses, the habitat recovers and the local survival probability increases until some maximal level is reached or the habitat is again disturbed. To capture the rapid decrease in the survival probability following disturbance, F_L should have considerable mass near one. The relatively slow recovery of the habitat means that F_R should have most of its mass near zero.

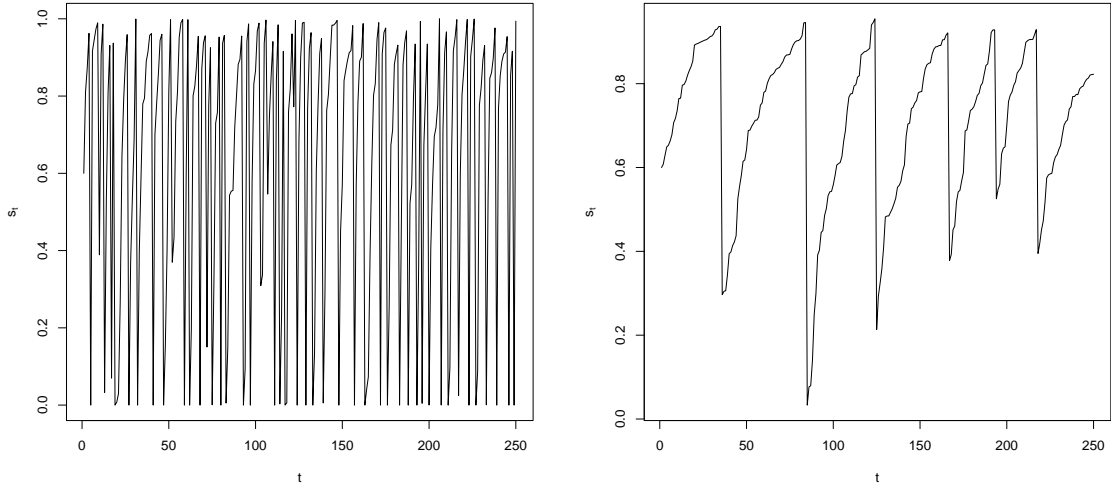


FIGURE 1. Simulated path of survival probabilities. Left: $L_t \sim \text{Beta}(1, 0.1)$, $R_t \sim \text{Beta}(1, 1)$. Right: $L_t \sim \text{Beta}(1, 1)$, $R_t \sim \text{Beta}(1, 20)$.

The function p reflects the probability of disturbance for a given survival probability and might reasonably be assumed to be increasing.

We simulate metapopulations with 50 and 250 habitat patches for 10^5 time steps with the two survival processes depicted in Figure 1 and compute the proportion of time each patch is occupied. Treating the resulting time series as stationary, the standard error on the estimated proportions was estimated to be no more than 0.003 for all simulations. This is compared to the fixed point of the deterministic recursion. Details of how the fixed point is calculated are given in Appendix C. The results are plotted in Figures 2. As we expect, the fixed point of the deterministic recursion provides a better approximation as the number of patches in the metapopulation increases. It appears that the deterministic recursion has a greater tendency to over-estimate the proportion of time the patch is occupied than to under-estimate it. Furthermore the deterministic recursion generally provides a better approximation for patches in the center of the metapopulation than those on the periphery.

5. DISCUSSION

The importance of landscape dynamics to the persistence of metapopulations has been well established in the literature. In contrast to previous mathematical analyses that have employed a suitable/unsuitable classification of habitat patches, we adopted a more

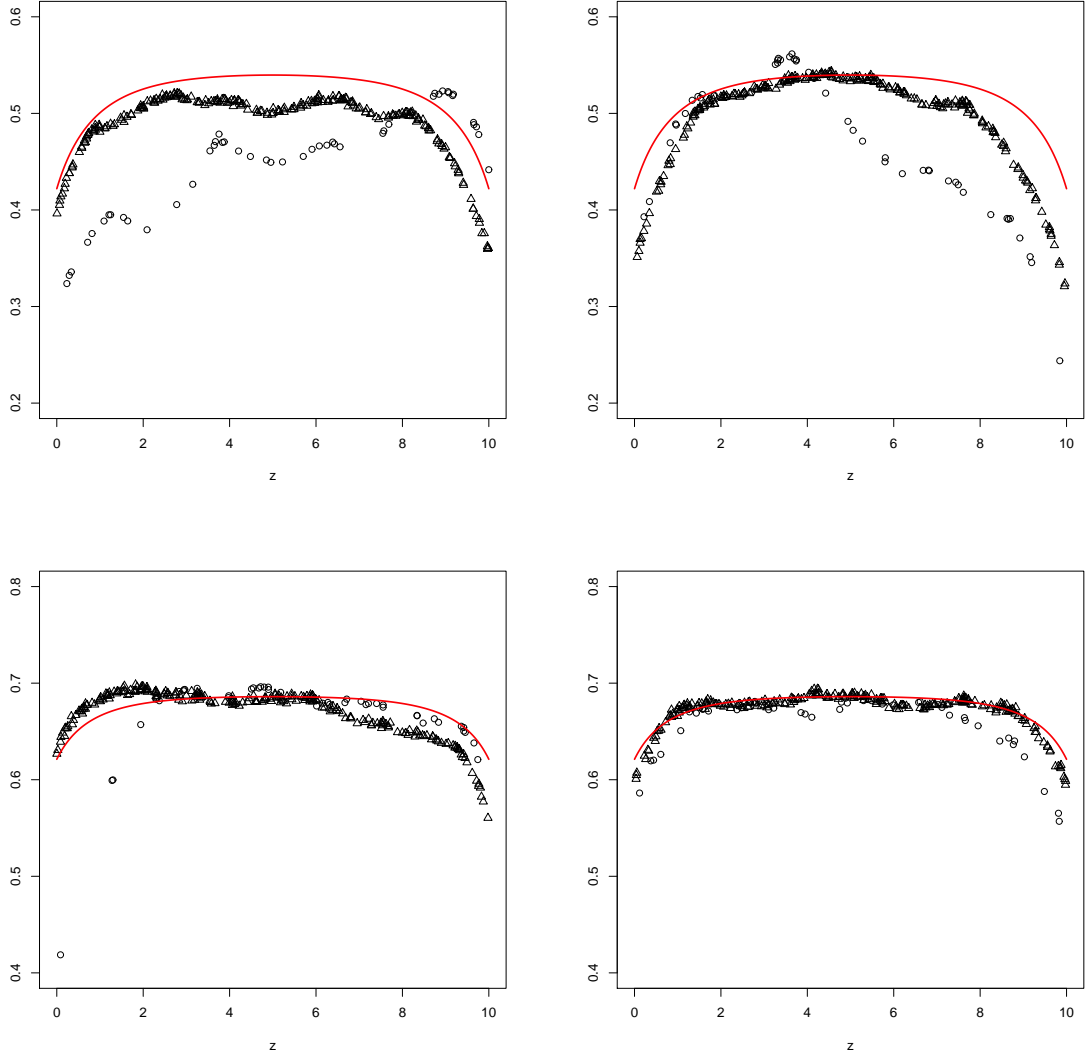


FIGURE 2. The estimated proportion of time each patch is occupied in a metapopulation with 50 patches (\circ) and 250 patches (\triangle) from simulations. The location of the patch is given by the position of the points on the z axis. The line indicates the limiting probability of patch occupancy. Landscape dynamics for top row took $L_t \sim \text{Beta}(1, 0.1)$ and $R_t \sim \text{Beta}(1, 1)$. Landscape dynamics for bottom row took $L_t \sim \text{Beta}(1, 1)$ and $R_t \sim \text{Beta}(1, 20)$.

general Markovian model for the landscape dynamics to better reflect environmental fluctuations. It would be of interest to combine the metapopulation model (2.4) with the Markovian models for succession studied in [60, 5, 35, 3] among others. The results

presented in Sections 3 and 4 would still apply since they were developed for general Markov landscape dynamics. However, using these specific landscape dynamics may reveal a more precise connection with metapopulation survival.

Our analysis yielded similar conclusions to those obtained from models employing the suitable/unsuitable classification of habitat patches. In particular, we note that when certain simplifications are imposed, our persistence criterion (Theorem 4.1) has a similar form to the persistence criterion for the spatially realistic Levins' model with dynamic landscape [14, 66]. As such, it seems that these conclusions are relatively robust to at least some modification of the model assumptions. Specifically, we believe the assumptions requiring the model to have the phase structure and for the patch characteristics to be identically distributed across space could be relaxed. Without the phase structure, we expect the existence and uniqueness of the equilibrium to still hold. In fact, much of the current proof could be retained with Lemma 8.2 being the main difficulty. On the other hand, to prove stability of the equilibrium would need different methods to those currently used. To relax the assumption that patch characteristics are identically distributed across space, we could allow the transition kernel of the patch characteristic, and hence its stationary distribution, to depend on the patch location. This would bring our model closer to the setting described in [14, 66] where differences in patch sizes and extinction rates are accommodated. Using the general results from Appendix A, we expect the connectivity measure to still have a deterministic limit and local populations at different patches to be asymptotically independent. However, to obtain a recursion similar to (3.9) - (3.11) which we needed to construct the persistence criterion in Theorem 4.1, would require a closer investigation of the reversed Markov chain for the patch characteristic.

Of the other assumptions used in the analysis, most are technical assumptions introduced to avoid certain pathological cases. The two assumptions which would significantly impact the results are that the colonisation function is concave and that fluctuations in the landscape are independent between habitat patches. As noted in Section 4, allowing non-concave colonisation functions, such as the one used in [22], would introduce the possibility of a strong Allee-like effect in the metapopulation [12] so the equilibrium would no longer be globally stable. Independence of patch characteristics at different patches has been used in many other metapopulation models incorporating landscape dynamics [31, 14, 55, 65, 66, 54]. However, for certain environmental disturbances such as fires,

droughts and floods, the spatial extent can be large compared to the entire habitable area [59], which means the assumption of independence between patches is unlikely to hold. If the metapopulation exhibited some limiting behaviour without the independence assumption, then it would most likely have a very different form.

Dependence between patch characteristics at different patches may not always be obvious. Theorem 3.2 offers the possibility of identifying dependence between patches when the metapopulation is large. For large metapopulations, if patch characteristics are independent at different patches, then the local populations at the two patches are approximately independent. We could estimate the strength of dependence between two patches. Strong dependence would indicate the presence of dependence in the patch characteristics. Unfortunately, testing for independence would not be very useful since the independence of local populations is only asymptotic.

One important way in which our model differs from those using the suitable/unsuitable classification of habitat patches is in the distribution of the local population life span. For static landscapes and dynamics landscapes using the suitable/unsuitable classification, the life span of a local population always has a geometric distribution (discrete time models) or an exponential distribution (continuous time models). We do not have any results characterising the life span distributions permitted by our model, but we expect that almost any distribution is possible by analogy with phase type distributions [47].

The effect of landscape dynamics on the equilibrium level of the metapopulation appears quite complicated in general, however for metapopulations with phase structure the role of landscape dynamics is much clearer. The landscape dynamics affect the equilibrium of the metapopulation primarily through the expected future contributions to the connectivity measure of a colonised patch. When the patch area is constant, expected future contributions to the connectivity measure of a colonised patch is determined by the distribution of the local population's life span. Given these results, it is natural to wonder whether the metapopulations with dynamic landscape behave similarly to metapopulations whose local populations have non-geometric/exponential life span distributions, at least in the specialised setting. However, in analysis not reported here we have seen that the equilibrium level in metapopulations with non-geometric life span distributions depends on the life span distribution only through its expectation. This is perhaps not surprising given this has also been observed in the SIS model with general infectious

period distributions [46]. (Recall the standard SIS model, or stochastic logistic model, has been used a stochastic counterpart to Levins' model [48].)

Finally, it has been observed in metapopulation models with suitable/unsuitable habitat dynamics that metapopulations are more likely to persist and will persist at higher levels of occupancy with static landscapes than with dynamic landscapes. Our analysis shows that this still holds for metapopulations with more general landscape dynamics (Corollary 4.4) if the patch area is constant. However, static landscapes are not necessarily optimal when the patch area is stochastic, leading to similar behaviour to pulsed dispersal [54]. The conclusion that static landscapes are optimal may also be false if we move beyond the metapopulation framework and consider species coexistence. Multiple mechanisms have been identified by which landscape dynamics enables the species coexistence [58, 56, 44]. These mechanisms are based on different species responding to environmental disturbances in different ways. For example, one species may be better at surviving disturbance, while another may actively try to colonise new areas in response to the disturbance. So although a static landscape may be optimal from the perspective of a single species, it may be to the detriment of other species in the community.

6. APPENDIX A — PROOFS FOR SINGLE PATCH ASYMPTOTICS

In this appendix we prove the results of Section 3 in a more general form than stated there. We begin by listing the main assumptions used in our analysis of model (2.4). For all n :

- (A) The functions $a(\cdot)$ and $s(\cdot)$ are continuous on Θ .
- (B) Both Θ and Ω are compact spaces.
- (C) The colonisation function is continuous in the second argument and satisfies the Lipschitz condition

$$\sup_{\theta \in \Theta} |f(z; \theta) - f(\tilde{z}; \theta)| \leq L |z - \tilde{z}|$$

for any $z, \tilde{z} \in \Omega$ and some $L < \infty$. The colonisation function is increasing and satisfies $f(0) = 0$.

- (D) The function $D(\cdot, \cdot)$ defines a uniformly bounded and equicontinuous family of functions on Ω . That is, there exists a finite constant \bar{D} such that for all $z, \tilde{z} \in \Omega$, $|D(z, \tilde{z})| \leq \bar{D}$, and for every $\epsilon > 0$ there exists a $\delta > 0$ such that for all z, \tilde{z} with

$$\|z - \tilde{z}\| < \delta$$

$$\sup_{y \in \Omega} |D(z, y) - D(\tilde{z}, y)| < \epsilon.$$

Furthermore, $D(z, \tilde{z}) > 0$ for all $(z, \tilde{z}) \in \Omega \times \Omega$.

(E) The transition kernel of the patch characteristic process satisfies the weak Feller property, that is, for every continuous function h on Θ , the function defined by

$$Ph(\theta) := \int_{\Theta} h(\eta) P(\theta, d\eta), \quad \theta \in \Theta,$$

is also continuous [43, Proposition 6.1.1(i)].

We will discuss these assumptions further in Subsection 6.3. For now we note that Assumptions (A) - (D) are satisfied by typical models. Assumption (E) is a regularity assumption needed when Θ is a general state space. It basically requires that the distributions $P(\theta, \cdot)$ and $P(\theta', \cdot)$ are close if θ and θ' are close. When Θ is a finite state space, Assumption (E) is trivially satisfied.

6.1. Limiting behaviour of the landscape. We construct random measures $\sigma_{n,t}$ which summarise the state of the landscape in a metapopulation with n patches at time t . These measures are purely atomic, placing mass n^{-1} at the point determined by patch i 's location and its characteristic variable at time t . Let $C^+(\Theta \times \Omega)$ be the space of continuous functions $h : \Theta \times \Omega \rightarrow [0, \infty)$. By Assumption (B), Ω and Θ are compact so every function in $C^+(\Theta \times \Omega)$ is bounded. The random measure $\sigma_{n,t}$ is defined by

$$\int_{\Theta \times \Omega} h(\theta, z) \sigma_{n,t}(d\theta, dz) := n^{-1} \sum_{i=1}^n h(\theta_{i,t}, z_i), \quad \text{for all } h \in C^+(\Theta \times \Omega).$$

As $n \rightarrow \infty$, the sequence of random measure $\sigma_{n,t}$ converges in distribution to σ_t if and only if

$$\int_{\Theta \times \Omega} h(\theta, z) \sigma_{n,t}(d\theta, dz) \xrightarrow{d} \int_{\Theta \times \Omega} h(\theta, z) \sigma_t(d\theta, dz), \quad \text{for all } h \in C^+(\Theta \times \Omega) \quad (6.21)$$

[29, Theorem 16.16]. Since we are only deal with random measures converging to non-random measures, the convergence in (6.21) can be replaced by convergence in probability. The last of our main assumptions is

(F) As $n \rightarrow \infty$, $\sigma_{n,0} \xrightarrow{d} \sigma_0$ for some non-random measure σ_0 .

Although this assumption only concerns the initial variation in the landscape, it implies a similar ‘law of large numbers’ for the landscape at all subsequent times.

Lemma 6.1. *Suppose Assumptions (B), (E) and (F) hold. Then $\sigma_{n,t} \xrightarrow{d} \sigma_t$, where σ_t is defined by the recursion*

$$\int_{\Theta \times \Omega} h(\theta, z) \sigma_{t+1}(d\theta, dz) = \int_{\Theta \times \Omega} h(\theta, z) \int_{\Omega} P(\eta, d\theta) \sigma_t(d\eta, dz), \quad \text{for all } h \in C^+(\Theta \times \Omega).$$

Proof. If $\int_{\Theta \times \Omega} h(\theta, z) \sigma_{n,t}(d\theta, dz) \xrightarrow{d} \int_{\Theta \times \Omega} h(\theta, z) \sigma_t(d\theta, dz)$ for all $h \in C^+(\Theta \times \Omega)$, then $\sigma_{n,t} \xrightarrow{d} \sigma_t$ [29, Theorem 16.16]. We use induction on t to prove weak convergence of the random measures $\sigma_{n,t}$ to non-random measures σ_t . By Assumption (F), $\sigma_{n,0} \xrightarrow{d} \sigma_0$ for some non-random measure σ_0 . The conditional expectation of $\int_{\Theta \times \Omega} h(\theta, z) \sigma_{n,t+1}(d\theta, dz)$ given (θ_t^n, z^n) is

$$\begin{aligned} \mathbb{E} \left(\int_{\Theta \times \Omega} h(\theta, z) \sigma_{n,t+1}(d\theta, dz) \mid \theta_t^n, z^n \right) &= n^{-1} \sum_{i=1}^n \int_{\Omega} h(\eta, z_i) P(\theta_{i,t}, d\eta) \\ &= \int_{\Theta \times \Omega} \left\{ \int_{\Omega} h(\eta, z) P(\theta, d\eta) \right\} \sigma_{n,t}(d\theta, dz). \end{aligned}$$

Suppose that $\sigma_{n,t} \xrightarrow{d} \sigma_t$ for some non-random measure σ_t . If $\int_{\Theta} h(\eta, z) P(\theta, d\eta)$ is in $C^+(\Theta \times \Omega)$, then

$$\lim_{n \rightarrow \infty} \mathbb{E} \left(\int_{\Theta \times \Omega} h(\theta, z) \sigma_{n,t+1}(d\theta, dz) \mid \theta_t^n, z^n \right) = \int_{\Theta \times \Omega} \left\{ \int_{\Omega} h(\eta, z) P(\theta, d\eta) \right\} \sigma_t(d\theta, dz). \quad (6.22)$$

We now show that $\int_{\Theta} h(\eta, z) P(\theta, d\eta) \in C^+(\Theta \times \Omega)$. For any $(\theta', z') \rightarrow (\theta, z)$,

$$\begin{aligned} \lim_{(\theta', z') \rightarrow (\theta, z)} \int_{\Theta} h(\eta, z') P(\theta', d\eta) &= \lim_{\theta' \rightarrow \theta} \int_{\Theta} h(\eta, z) P(\theta', d\eta) \\ &\quad + \lim_{(\theta', z') \rightarrow (\theta, z)} \int_{\Theta} [h(\eta, z') - h(\eta, z)] P(\theta', d\eta). \end{aligned}$$

Since P has the weak Feller property from Assumption (E),

$$\lim_{\theta' \rightarrow \theta} \int_{\Theta} h(\eta, z) P(\theta', d\eta) = \int_{\Theta} h(\eta, z) P(\theta, d\eta).$$

As $\int_{\Theta} P(\theta, d\eta) = 1$,

$$\left| \lim_{(\theta', z') \rightarrow (\theta, z)} \int_{\Theta} [h(\eta, z') - h(\eta, z)] P(\theta', d\eta) \right| \leq \lim_{z' \rightarrow z} \sup_{\theta' \in \Theta} |h(\theta', z') - h(\theta', z)|.$$

From Assumption (B), $\Theta \times \Omega$ is compact so the Heine-Cantor Theorem implies that h is uniformly continuous. Therefore, $\int_{\Theta} h(\eta, z') P(\theta', d\eta) \rightarrow \int_{\Theta} h(\eta, z) P(\theta, d\eta)$ as $(\theta', z') \rightarrow (\theta, z)$. Hence, $\int_{\Theta} h(\eta, z) P(\theta, d\eta) \in C^+(\Theta \times \Omega)$ and equality (6.22) holds.

The conditional variance of $\int_{\Theta \times \Omega} h(\theta, z) \sigma_{n,t+1}(d\theta, dz)$ can be bounded by

$$\text{var} \left(\int_{\Theta \times \Omega} h(\theta, z) \sigma_{n,t+1}(d\theta, dz) \mid \theta_t^n, z_t^n \right) \leq n^{-1} \sup_{(\theta, z) \in \Theta \times \Omega} |h(\theta, z)|^2.$$

As the conditional variance goes to zero in probability, we can apply a Chebyshev type inequality [38, Appendix C] to conclude that

$$\begin{aligned} \int_{\Theta \times \Omega} h(\theta, z) \sigma_{n,t+1}(d\theta, dz) &\xrightarrow{p} \int_{\Theta \times \Omega} \left\{ \int_{\Omega} h(\eta, z) P(\theta, d\eta) \right\} \sigma_t(d\theta, dz). \\ &= \int_{\Theta \times \Omega} h(\theta, z) \left\{ \int_{\Omega} P(\eta, d\theta) \sigma_t(d\eta, dz) \right\} \\ &= \int_{\Theta \times \Omega} h(\theta, z) \sigma_{t+1}(d\theta, dz). \end{aligned} \quad (6.23)$$

Hence, $\sigma_{n,t+1} \xrightarrow{d} \sigma_{t+1}$. The recursion for σ_{t+1} is determined by equation (6.23). \square

6.2. Limiting behaviour of the metapopulation. Similar to our treatment of the landscape, we construct random measures $\mu_{n,t}$ which summarise the state of the metapopulation at time t . These measures are defined by

$$\int_{\Theta \times \Omega} h(\theta, z) \mu_{n,t}(d\theta, dz) := n^{-1} \sum_{i=1}^n X_{i,t}^n h(\theta_{i,t}, z_i), \quad \text{for all } h \in C^+(\Theta \times \Omega).$$

The measure $\mu_{n,t}$ has a similar structure to $\sigma_{n,t}$, but only involves those patches that are occupied at time t . Under the stated assumptions, the sequence of random measures $\{\mu_{n,t}\}_{n=1}^{\infty}$ converges to a deterministic measure as the number of patches tends to infinity.

Theorem 6.2. *Suppose that Assumptions (A) – (F) hold and that $\mu_{n,0} \xrightarrow{d} \mu_0$ for some non-random measure μ_0 . Then $\mu_{n,t} \xrightarrow{d} \mu_t$ for all $t = 0, 1, \dots$, where μ_t is defined by the recursion*

$$\begin{aligned} &\int_{\Theta \times \Omega} h(\theta, z) \mu_{t+1}(d\theta, dz) \\ &= \int_{\Theta \times \Omega} s(\theta) \left\{ \int_{\Theta} h(\eta, z) P(\theta, d\eta) \right\} \mu_t(d\theta, dz) \\ &\quad + \int_{\Theta \times \Omega} \left\{ \int_{\Theta} h(\eta, z) P(\theta, d\eta) \right\} f \left(\int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z}); \theta \right) \sigma_t(d\theta, dz) \\ &\quad - \int_{\Theta \times \Omega} \left\{ \int_{\Theta} h(\eta, z) P(\theta, d\eta) \right\} f \left(\int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z}); \theta \right) \mu_t(d\theta, dz), \end{aligned} \quad (6.24)$$

for all $h \in C^+(\Theta \times \Omega)$.

Proof. The proof follows closely the arguments of the proof of Lemma 6.1 and the proof of Theorem 3.1 [41]. By assumption $\mu_{n,0} \xrightarrow{d} \mu_0$ for some non-random measure μ_0 . Suppose that $\mu_{n,t} \xrightarrow{d} \mu_t$ for some non-random measure μ_t . Then

$$\begin{aligned} & \mathbb{E} \left(\int_{\Theta \times \Omega} h(\theta, z) \mu_{n,t+1}(d\theta, dz) \mid X_t^n, \theta_t^n, z^n \right) \\ &= n^{-1} \sum_{i=1}^n \mathbb{E} (h(\theta_{i,t+1}, z_i) \mid \theta_{i,t}, z_i) \mathbb{E} (X_{i,t+1}^n \mid X_t^n, \theta_t^n, z^n) \\ &= \int_{\Theta \times \Omega} s(\theta) \left\{ \int_{\Theta} h(\eta, z) P(\theta, d\eta) \right\} \mu_{n,t}(d\theta, dz) \end{aligned} \quad (6.25)$$

$$+ \int_{\Theta \times \Omega} \left\{ \int_{\Theta} h(\eta, z) P(\theta, d\eta) \right\} f \left(\int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z}); \theta \right) \sigma_{n,t}(d\theta, dz) \quad (6.26)$$

$$- \int_{\Theta \times \Omega} \left\{ \int_{\Theta} h(\eta, z) P(\theta, d\eta) \right\} f \left(\int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z}); \theta \right) \mu_{n,t}(d\theta, dz) + \epsilon_{n,t}(h), \quad (6.27)$$

where

$$\begin{aligned} |\epsilon_{n,t}(h)| &\leq 2L \left(\int_{\Theta \times \Omega} \int_{\Theta} h(\eta, z) P(\theta, d\eta) \sigma_n(d\theta, dz) \right) \\ &\quad \times \sup_{z \in \Omega} \left| \int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_{n,t}(d\tilde{\theta}, d\tilde{z}) - \int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z}) \right|, \end{aligned}$$

as f is uniformly Lipschitz continuous from Assumption (C). Ranga Rao [52, Theorem 3.1] showed that

$$\sup_{g \in \mathcal{G}} \left| \int_{\mathcal{X}} g(x) \nu_n(dx) - \int_{\mathcal{X}} g(x) \nu(dx) \right| \rightarrow 0,$$

for a sequence of probability measures ν_n converging weakly to ν and \mathcal{G} a uniformly bounded and equicontinuous family of functions on \mathcal{X} . Applying a small modification that result and Assumption (D), it follows that if $\mu_{n,t} \xrightarrow{d} \mu_t$, a non-random measure, then

$$\sup_{z \in \Omega} \left| \int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_{n,t}(d\tilde{\theta}, d\tilde{z}) - \int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z}) \right| \xrightarrow{p} 0.$$

To prove convergence of the integrals at (6.25) - (6.27), we need both $s(\theta) \int_{\Theta} h(\eta, z) P(\theta, d\eta)$ and $\int_{\Theta} h(\eta, z) P(\theta, d\eta) f(\int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z}); \theta)$ to be in $C^+(\Theta \times \Omega)$. From the proof of Lemma 6.1, $\int_{\Theta} h(\eta, z) P(\theta, d\eta) \in C^+(\Theta \times \Omega)$. With Assumption (A) this implies $s(\theta) \int_{\Theta} h(\eta, z) P(\theta, d\eta) \in C^+(\Theta \times \Omega)$. Also $f(\int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z}); \theta) \in C^+(\Theta \times \Omega)$ by Assumptions (C) and (D) so $\int_{\Theta} h(\eta, z) P(\theta, d\eta) f(\int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z}); \theta) \in$

$C^+(\Theta \times \Omega)$. Applying the induction hypothesis, $\mu_{n,t} \xrightarrow{d} \mu_t$ for some non-random measure μ_t . Therefore,

$$\begin{aligned} & \mathbb{E} \left(\int_{\Theta \times \Omega} h(\theta, z) \mu_{n,t+1}(d\theta, dz) \mid X_t^n, \theta_t^n, z_t^n \right) \xrightarrow{p} \int_{\Theta \times \Omega} s(\theta) \left\{ \int_{\Theta} h(\eta, z) P(\theta, d\eta) \right\} \mu_t(d\theta, dz) \\ & + \int_{\Theta \times \Omega} \left\{ \int_{\Theta} h(\eta, z) P(\theta, d\eta) \right\} f \left(\int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z}); \theta \right) \sigma_t(d\theta, dz) \\ & - \int_{\Theta \times \Omega} \left\{ \int_{\Theta} h(\eta, z) P(\theta, d\eta) \right\} f \left(\int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z}); \theta \right) \mu_t(d\theta, dz). \end{aligned}$$

The conditional variance of $\int_{\Theta \times \Omega} h(\theta, z) \mu_{n,t+1}(d\theta, dz)$ can be bounded by $n^{-1} \sup_{(\theta, z)} |h(\theta, z)|^2$.

Applying a Chebyshev type inequality [38, Appendix C], we conclude that $\int_{\Theta \times \Omega} h(\theta, z) \mu_{n,t+1}(d\theta, dz)$ converges to $\int_{\Theta \times \Omega} h(\theta, z) \mu_{t+1}(d\theta, dz)$ in probability. Hence, $\mu_{n,t+1} \xrightarrow{d} \mu_{t+1}$ with μ_{t+1} determined by the recursion (6.24). \square

A consequence of Theorem 6.2 is that $(X_{i,t}^n, \theta_{i,t})$ converges to a Markov chain with time dependent transition probabilities.

Corollary 6.3. *Assume the conditions of Theorem 6.2 hold. If $X_{i,0}^n \xrightarrow{p} X_{i,0}$, then $X_{i,t}^n \xrightarrow{p} X_{i,t}$ for all $t \geq 0$, where the transition probability for $X_{i,t}$ is*

$$\mathbb{P}(X_{i,t+1} = 1 \mid X_{i,t}, \theta_{i,t} = \theta, z_i) = s(\theta) X_{i,t} + f(\psi_t(z_i); \theta) (1 - X_{i,t}) \quad (6.28)$$

and

$$\psi_t(z) = \int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z}) \quad (6.29)$$

Proof. The proof follows the same arguments as the proof of Corollary 1 of McVinish and Pollett [39]. \square

The following result assumes that the landscape is in equilibrium to simplify the recursion (6.24).

Theorem 6.4. *If $\sigma_t = \sigma$ for some measure σ and all $t \geq 0$, then μ_t is absolutely continuous with respect to σ for all $t \geq 0$. The Radon-Nikodým derivative of μ_t with*

respect σ , denoted by $\frac{\partial \mu_t}{\partial \sigma}$, is given by the recursion

$$\begin{aligned} \frac{\partial \mu_{t+1}}{\partial \sigma}(\theta, z) &= \int_{\Theta} s(\eta) \frac{\partial \mu_t}{\partial \sigma}(\eta, z) P^*(\theta, d\eta) \\ &\quad + \int_{\Theta} f \left(\int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \frac{\partial \mu_t}{\partial \sigma}(\tilde{\theta}, \tilde{z}) \sigma(d\tilde{\theta}, d\tilde{z}); \eta \right) \left(1 - \frac{\partial \mu_t}{\partial \sigma}(\eta, z) \right) P^*(\theta, d\eta). \end{aligned} \quad (6.30)$$

Proof. That μ_t is absolutely continuous with respect to σ for all $t \geq 0$ follows from the same arguments as McVinish and Pollett [40, Lemma 5]. The recursion for the Radon-Nikodým derivative of μ_t with respect to σ requires the dual kernel (see Appendix B). Applying Corollary 7.2 to the integrals on the right hand side of recursion (6.24), we can express the three terms as

$$\begin{aligned} &\int_{\Theta \times \Omega} s(\theta) \left\{ \int_{\Theta} h(\eta, z) P(\theta, d\eta) \right\} \mu_t(d\theta, dz) \\ &= \int_{\Theta} h(\theta, z) \left\{ \int_{\Theta \times \Omega} s(\eta) \frac{\partial \mu_t}{\partial \sigma}(\eta, z) P^*(\theta, d\eta) \right\} \sigma(d\eta, dz), \\ &\int_{\Theta \times \Omega} \left\{ \int_{\Theta} h(\eta, z) P(\theta, d\eta) \right\} f \left(\int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z}); \theta \right) \sigma(d\theta, dz) \\ &= \int_{\Theta \times \Omega} h(s, z) \left\{ \int_{\Theta} f \left(\int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z}); \eta \right) P^*(\theta, d\eta) \right\} \sigma(d\theta, dz), \end{aligned}$$

and

$$\begin{aligned} &\int_{\Theta \times \Omega} \left\{ \int_{\Theta} h(\eta, z) P(\theta, d\eta) \right\} f \left(\int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z}); \theta \right) \mu_t(d\theta, dz) \\ &= \int_{\Theta \times \Omega} h(s, z) \left\{ \int_{\Theta} f \left(\int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z}); \eta \right) \frac{\partial \mu_t}{\partial \sigma}(\eta, z) P^*(\theta, d\eta) \right\} \sigma(d\theta, dz). \end{aligned}$$

The recursion follows by combining these terms and noting that the Radon-Nikodým derivative is uniquely defined up to a σ -null set. \square

Theorem 6.5. Assume that $(X_{i,0}^n, \theta_{i,0}, z_i) \stackrel{d}{=} (X_{j,0}^n, \theta_{j,0}, z_j)$ for all i, j . Suppose that

$$\mathbb{P}(X_{i,0}^n = 1 \mid \theta_{i,0} = \theta, z_i = z) = q_0(\theta, z),$$

for some function $q_0 : \Theta \times \Omega \rightarrow [0, 1]$. Under the conditions of Corollary 6.3 and Theorem 6.4,

$$\mathbb{P}(X_{i,t} = 1 \mid \theta_{i,t} = \theta, z_i = z) = \frac{\partial \mu_t}{\partial \sigma}(\theta, z) \quad (6.31)$$

for all $t \geq 0$.

Proof. We first show that σ is the distribution of $(\theta_{i,0}, z_i)$. For any $h \in C^+(\Theta \times \Omega)$, the sequence of random variables $\int_{\Theta \times \Omega} h(\theta, z) \sigma_{n,0}(d\theta, dz)$ is uniformly integrable. This sequence of random variables converges in probability as the limiting measure σ is non-random. Therefore,

$$\mathbb{E} \left(\int_{\Theta \times \Omega} h(\theta, z) \sigma_{n,0}(d\theta, dz) \right) \rightarrow \int_{\Theta \times \Omega} h(\theta, z) \sigma(d\theta, dz). \quad (6.32)$$

As $(\theta_{i,0}, z_i) \stackrel{d}{=} (\theta_{j,0}, z_j)$ for all i, j ,

$$\mathbb{E} \left(\int_{\Theta \times \Omega} h(\theta, z) \sigma_{n,0}(d\theta, dz) \right) = \mathbb{E} (h(\theta_{i,0}, z_i)). \quad (6.33)$$

Since (6.32) and (6.33) hold for all $h \in C^+(\Theta \times \Omega)$, we see that σ is the distribution of $(\theta_{i,0}, z_i)$.

Similarly, we can identify the Radon-Nikodým derivative of μ_0 with respect to σ as the function $q_0(\theta, z)$. For any $h \in C^+(\Theta \times \Omega)$, the sequence of random variables $\int_{\Theta \times \Omega} h(\theta, z) \mu_{n,0}(d\theta, dz)$ converges in probability and is uniformly integrable. Therefore,

$$\mathbb{E} \left(\int_{\Theta \times \Omega} h(\theta, z) \mu_{n,0}(d\theta, dz) \right) \rightarrow \int_{\Theta \times \Omega} h(\theta, z) \mu_0(d\theta, dz) = \int_{\Theta \times \Omega} h(\theta, z) \frac{\partial \mu_0}{\partial \sigma}(\theta, z) \sigma(d\theta, dz).$$

As $(X_{i,0}^n, \theta_{i,0}, z_i) \stackrel{d}{=} (X_{j,0}^n, \theta_{j,0}, z_j)$ for all i, j ,

$$\begin{aligned} \mathbb{E} \left(\int_{\Theta \times \Omega} h(\theta, z) \mu_{n,0}(d\theta, dz) \right) &= \mathbb{E} (h(\theta_{i,0}, z_i) X_{i,0}^n) \\ &= \int_{\Theta \times \Omega} h(\theta, z) \mathbb{P} (X_{i,0}^n = 1 \mid \theta_{i,0} = \theta, z_i = z) \sigma(d\theta, dz) \\ &= \int_{\Theta \times \Omega} h(\theta, z) q_0(\theta, z) \sigma(d\theta, dz). \end{aligned} \quad (6.34)$$

As equation (6.34) holds for all $h \in C^+(\Theta \times \Omega)$ and the Radon-Nikodým derivative is unique, it follows that

$$\frac{\partial \mu_0}{\partial \sigma}(\theta, z) = q_0(\theta, z).$$

Thus, we have established equality (6.31) for $t = 0$. The proof for $t > 0$ proceeds by induction. Although we will always be conditioning on the patch location, this will not be made explicit to simplify the expressions. Let $\psi_t(z) = \int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mu_t(d\tilde{\theta}, d\tilde{z})$. Then

$(X_{i,t}, \theta_{i,t})$ is a Markov chain on $\{0, 1\} \times \Theta$ with transition kernel

$$\begin{aligned}\mathbb{P}(X_{i,t+1} = 1, \theta_{i,t+1} \in A \mid X_{i,t} = x, \theta_{i,t} = \theta) &= (s(\theta)x + f(\psi_t(z); \theta)(1-x)) \int_A P(\theta, d\eta) \\ \mathbb{P}(X_{i,t+1} = 0, \theta_{i,t+1} \in A \mid X_{i,t} = x, \theta_{i,t} = \theta) &= ((1-s(\theta))x + (1-f(\psi_t(z); \theta))(1-x)) \int_A P(\theta, d\eta),\end{aligned}$$

for any measurable set $A \subset \Theta$. To compute $\mathbb{P}(X_{i,t} = 1 \mid \theta_{i,t} = \theta)$, note that

$$\begin{aligned}\mathbb{P}(X_{i,t+1} = 1, \theta_{i,t+1} \in A) & \tag{6.35} \\ &= \int_{\Theta \times \Omega} \mathbb{P}(X_{i,t+1} = 1, \theta_{i,t+1} \in A \mid X_{i,t} = 1, \theta_{i,t} = \theta) \mathbb{P}(X_{i,t} = 1 \mid \theta_{i,t} = \theta) \sigma(d\theta, dz) \\ &+ \int_{\Theta \times \Omega} \mathbb{P}(X_{i,t+1} = 1, \theta_{i,t+1} \in A \mid X_{i,t} = 0, \theta_{i,t} = \theta) \mathbb{P}(X_{i,t} = 0 \mid \theta_{i,t} = \theta) \sigma(d\theta, dz) \\ &= \int_{\Theta \times \Omega} \left\{ \int_A P(\theta, d\eta) \right\} s(\theta) \mathbb{P}(X_{i,t} = 1 \mid \theta_{i,t} = \theta) \sigma(d\theta, dz) \\ &+ \int_{\Theta \times \Omega} \left\{ \int_A P(\theta, d\eta) \right\} f(\psi_t(z); \theta) (1 - \mathbb{P}(X_{i,t} = 1 \mid \theta_{i,t} = \theta)) \sigma(d\theta, dz). \tag{6.36}\end{aligned}$$

Applying Corollary 7.2 to the integrals in (6.36) gives

$$\begin{aligned}& \int_{\Theta \times \Omega} \int_{\Theta} \mathbf{1}(\eta \in A) P(\theta, d\eta) s(\theta) \mathbb{P}(X_{i,t} = 1 \mid \theta_{i,t} = \theta) \sigma(d\theta, dz) \\ &= \int_{\Theta \times \Omega} \left\{ \int_{\Theta} P^*(\theta, d\eta) s(\eta) \mathbb{P}(X_{i,t} = 1 \mid \theta_{i,t} = \eta) \right\} \mathbf{1}(\theta \in A) \sigma(d\theta, dz) \tag{6.37}\end{aligned}$$

and

$$\begin{aligned}& \int_{\Theta \times \Omega} \int_{\Theta} \mathbf{1}(\eta \in A) P(\theta, d\eta) f(\psi_t(z); \theta) (1 - \mathbb{P}(X_{i,t} = 1 \mid \theta_{i,t} = \theta)) \sigma(d\theta, dz) \\ &= \int_{\Theta \times \Omega} \left\{ \int_{\Theta} P^*(\theta, d\eta) f(\psi_t(z); \eta) (1 - \mathbb{P}(X_{i,t} = 1 \mid \theta_{i,t} = \eta)) \right\} \mathbf{1}(\theta \in A) \sigma(d\theta, dz). \tag{6.38}\end{aligned}$$

Substituting (6.37) and (6.38) into equation (6.36) yields

$$\begin{aligned}\mathbb{P}(X_{i,t+1} = 1, \theta_{i,t+1} \in A) & \\ &= \int_{\Theta \times \Omega} \left\{ \int_{\Theta} P^*(\theta, d\eta) s(\eta) \mathbb{P}(X_{i,t} = 1 \mid \theta_{i,t} = \eta) \right\} \mathbf{1}(\theta \in A) \sigma(d\theta, dz) \\ &+ \int_{\Theta \times \Omega} \left\{ \int_{\Theta} f(\psi_t(z); \eta) P^*(\theta, d\eta) s(\eta) (1 - \mathbb{P}(X_{i,t} = 1 \mid \theta_{i,t} = \eta)) \right\} \mathbf{1}(\theta \in A) \sigma(d\theta, dz).\end{aligned}$$

As the Radon-Nikodým derivative is unique up to a σ -null set,

$$\begin{aligned}\mathbb{P}(X_{i,t+1} = 1 \mid \theta_{i,t+1} = \theta) & \\ &= \int_{\Theta} s(\eta) \mathbb{P}(X_{i,t} = 1 \mid \theta_{i,t} = \eta) P^*(\theta, d\eta) + \int_{\Theta} f(\psi_t(z); \eta) (1 - \mathbb{P}(X_{i,t} = 1 \mid \theta_{i,t} = \eta)) P^*(\theta, d\eta).\end{aligned}$$

Comparing with (6.30), we see that if

$$\mathbb{P}(X_{i,t} = 1 \mid \theta_{i,t} = \theta) = \frac{\partial \mu_t}{\partial \sigma}(\theta, z),$$

then

$$\mathbb{P}(X_{i,t+1} = 1 \mid \theta_{i,t+1} = \theta) = \frac{\partial \mu_{t+1}}{\partial \sigma}(\theta, z),$$

for all $t \geq 0$. □

6.3. Proof of results from Section 3.

Proof of Lemma 3.1. For any $h \in C(\Theta \times \Omega)$, $\mathbb{E}(h(\theta_{i,t}, z_i)) = \mathbb{E}(\mathbb{E}(h(\theta_{i,t}, z_i) \mid \theta_{i,0}, z_i))$. Let P^t be the t -step transition kernel of the Markov chain for patch characteristic. As this Markov chain is positive Harris and aperiodic, it has a unique invariant measure π and, by Meyn and Tweedie [43, Theorem 13.3.3],

$$\mathbb{E}(h(\theta_{i,t}, z_i) \mid \theta_{i,0} = \theta, z_i = z) = \int_{\Theta} h(\eta, z) P^t(\theta, d\eta) \rightarrow \int_{\Theta} h(\eta, z) \pi(d\eta),$$

for every (θ, z) as $t \rightarrow \infty$. By the Dominated Convergence Theorem,

$$\lim_{t \rightarrow \infty} \mathbb{E}(h(\theta_{i,t}, z_i)) = \mathbb{E}\left(\int_{\Theta} h(\eta, z_i) \pi(d\eta)\right) = \int_{\Omega} \int_{\Theta} h(\eta, z) \pi(d\eta) \zeta(z) dz. \quad (6.39)$$

As (6.39) holds for all $h \in C(\Theta \times \Omega)$, this proves the limit (3.7) □

The results given in Section 3 follow as special cases of the results in this appendix. The assumptions required for these results to hold are rather weak. Assumption (B) requires Ω to be compact, that is a closed and bounded subset of \mathbb{R}^d . The condition imposed by Assumption (B) on Θ is trivially satisfied when Θ is a finite set, so too Assumptions (A) and (E). Assumption (C) will be satisfied when Θ is a finite set if for each $\theta \in \Theta$, the colonisation function $f(\cdot, \theta)$ is Lipschitz. This is a fairly weak requirement since in most cases the colonisation function is taken to be smooth. Even for general state spaces, Assumptions (A), (B) and (E) are not very restrictive. For example, if the Markov chain for the patch characteristic can be expressed as $\theta_t = F(\theta_{t-1}, W_t)$ with the W_t independent and identically distributed random variables taking values in \mathbb{R}^m and $F : \Theta \times \mathbb{R}^m \rightarrow \Theta$ a smooth function, then Assumption (E) is satisfied. Similar to the finite state space case, Assumption (C) will be satisfied for a general state space if for each $\theta \in \Theta$, the colonisation function $f(\cdot, \theta)$ is Lipschitz and the Lipschitz constant is bound on Θ . Finally, Assumption (D) is satisfied by the typical choice $D(z, z') = \exp(-\alpha \|z - z'\|)$, but range

limited dispersion kernels such as $D(z, z') \propto \mathbf{1}(\|z - z'\| \leq r)$ will not satisfy the lower bound imposed in Assumption (D).

Assume now that the $(X_{i,0}^n, \theta_{i,0}, z_i)$ are independent and identically distributed random variables. By the law of large numbers and Lemma 3.1,

$$\int_{\Theta \times \Omega} h(\theta, z) \sigma_{n,0}(d\theta, dz) \xrightarrow{d} \int_{\Omega} \int_{\Theta} h(\theta, z) \pi(d\theta) \zeta(z) dz, \quad \text{for all } h \in C^+(\Theta \times \Omega).$$

Therefore, $\sigma_{n,0}$ converges in distribution to the non-random measure $\pi \times \zeta$ and Assumption (F) holds. Furthermore, $\sigma_t = \pi \times \zeta$ as $(\theta_{i,t}, z_i) \stackrel{d}{=} (\theta_{i,0}, z_i)$ for all i . The law of large numbers can also be used to show $\mu_{n,0} \xrightarrow{d} \mu_0$. As in the proof of Theorem 6.5, if equation (3.8) holds, then q_0 is the Random-Nikodým derivative of μ_0 with respect to σ . Theorem 3.2 now follows from Corollary 6.3 and Theorem 6.4. Theorem 3.3 follows from Theorem 6.5 and Theorem 3.4 follows from Theorem 6.2.

7. APPENDIX B — DUAL PROCESS CONSTRUCTION

The dual kernel has been used by various authors studying Markov chains and processes [see 7, and references therein]. As we have been unable to find anything in the literature dealing explicitly with the case of interest here, we state the definition of the dual kernel and some basic results. In the following, (S, Σ) denotes a general measurable space. If $P = \{P(x, A), x \in S, A \in \Sigma\}$ is such that (i) for each $A \in \Sigma$, $P(\cdot, A)$ is a non-negative measurable function on S , and (ii) for each $x \in S$, $P(x, \cdot)$ is a measure on (S, Σ) with $P(x, S) \leq 1$, then we call P a sub-transition kernel. If $P(x, S) = 1$ for all $x \in S$, then P is a transition kernel [43, pg. 65]

Definition 7.1. Let P be a sub-transition kernel on (S, Σ) and let π be a σ -finite measure on (S, Σ) . If there exists a sub-transition kernel P^* such that

$$\int_A \pi(dx) P(x, B) = \int_B \pi(dx) P^*(x, A), \quad (7.40)$$

for all $A, B \in \Sigma$, then P^* is called a *dual of P with respect to π* . If

$$\int_A \pi(dx) P(x, B) = \int_B \pi(dx) P(x, A), \quad (7.41)$$

for all $A, B \in \Sigma$, then P is said to be *reversible with respect to π* .

We shall see that if π is a subinvariant measure for P , then the dual of P with respect to π is determined uniquely π -almost everywhere, in that, for all $A \in \Sigma$, $P^*(x, A)$ is the

same for π -almost all $x \in S$. By setting B equal to S in (7.41), we notice that if P is reversible with respect to π , then π is an invariant measure for P . More generally, we have the following.

Theorem 7.1. *Let P be a sub-transition kernel on (S, Σ) and let π be a σ -finite measure on (S, Σ) . Then π is a subinvariant measure for P if and only if there exists a dual P^* for P with respect to π . Further, π is an invariant measure for P if and only if P^* is a transition kernel. If P^* is dual for P , then π is invariant for P^* if and only if P is a transition kernel.*

Proof. Let P be a sub-transition kernel on (S, Σ) and let π be a σ -finite measure on (S, Σ) . We first show that if π is a subinvariant measure for P , then there exists a sub-transition kernel P^* satisfying Definition 7.1. Suppose π is subinvariant for P . For $A \in \Sigma$, define

$$\eta_A(\cdot) := \int_A \pi(dx) P(x, \cdot).$$

It is a measure on (S, Σ) because $P(x, \cdot)$ is a measure on (S, Σ) . It is also clear that η_A is absolutely continuous with respect to π , because if $N \in \Sigma$ is any π -null set then

$$\eta_A(N) = \int_A \pi(dx) P(x, N) \leq \int_S \pi(dx) P(x, N) \leq \pi(N) = 0.$$

So, by the Radon-Nikodým theorem, there exists a function $P^* : S \times \Sigma \rightarrow [0, \infty)$ such that $P^*(\cdot, A)$ is a Σ -measurable function, and for all $B \in \Sigma$,

$$\int_A \pi(dx) P(x, B) = \eta_A(B) = \int_B \pi(dx) P^*(x, A).$$

Hence, P^* is determined uniquely π -almost everywhere by equation (7.40). It remains to show that, for π -almost all $x \in S$, $P^*(x, \cdot)$ is a measure on (S, Σ) with $P^*(x, S) \leq 1$.

For any $A \in \Sigma$, $P^*(\cdot, A)$ is the Radon-Nikodým derivative of η_A with respect to π . As η_\emptyset is the null measure, $P^*(x, \emptyset) = 0$ for π -almost all $x \in S$. To show that $P^*(x, \cdot)$ is countably additive, let $\{B_k\}$ be a sequence of pairwise disjoint sets in Σ . We want to show that the Radon-Nikodým derivative of $\eta_{\cup_k B_k}$ with respect to π is $\sum_k P^*(\cdot, B_k)$. For

any $A \in \Sigma$,

$$\begin{aligned}
\eta_{\cup_k B_k}(A) &= \int_{\cup_k B_k} \pi(dx) P(x, A) \\
&= \sum_k \int_{B_k} \pi(dx) P(x, A) \\
&= \sum_k \int_A \pi(dx) P^*(x, B_k) \\
&= \int_A \pi(dx) \sum_k P^*(x, B_k).
\end{aligned}$$

Hence, $P^*(x, \cup_k B_k) = \sum_k P^*(x, B_k)$ for π -almost all $x \in S$. Finally, since π is subinvariant for P , we have, for any $A \in \Sigma$,

$$\int_A \pi(dx) P^*(x, S) = \int_S \pi(dx) P(x, A) \leq \pi(A).$$

Hence, by the Radon-Nikodým Theorem, $P^*(x, S) \leq 1$ for π -almost all $x \in S$.

We now show that if there exists a dual P^* for P with respect to π , then π is subinvariant. Since P^* is a sub-transition kernel, $P^*(x, S) \leq 1$ for all $x \in S$. On setting B equal to S in equation (7.40) we see that

$$\int_S \pi(dx) P(x, A) = \int_A \pi(dx) P^*(x, S) \leq \int_A \pi(dx) = \pi(A), \quad (7.42)$$

that is, π is subinvariant for P . This completes the proof of the first part of Theorem 7.1.

To prove the second part we note that if P^* is a transition kernel then $P^*(x, S) = 1$ for all $x \in S$. In that case, inequality (7.42) becomes equality, and π is seen to be invariant. On the other hand, if π is invariant for P , then

$$\pi(A) = \int_S \pi(dx) P(x, A) = \int_A \pi(dx) P^*(x, S),$$

for all $A \in \Sigma$. Therefore, $P^*(x, S) = 1$ for π -almost all $x \in S$, and P^* is a transition kernel. The final part is proved in similar vein. \square

Corollary 7.2. *Let ϕ and ψ be Σ -measurable functions. Then, under the conditions of Theorem 7.1, the dual P^* satisfies*

$$\int_S \pi(dx) \phi(x) \int_S P(x, dy) \psi(y) = \int_S \pi(dx) \psi(x) \int_S P^*(x, dy) \phi(y).$$

Proof. Suppose the conditions of Theorem 7.1 hold, and P^* is a sub-transition kernel that satisfies equation (7.40). Let ϕ and ψ be the indicator functions $\phi(x) = \mathbb{I}(x \in A)$ and $\psi(x) = \mathbb{I}(x \in B)$, where A, B are Σ -measurable sets. Then

$$\begin{aligned} \int_S \pi(dx) \phi(x) \int_S P(x, dy) \psi(y) &= \int_S \pi(dx) \mathbb{I}(x \in A) \int_S P(x, dy) \mathbb{I}(y \in B) \\ &= \int_A \pi(dx) P(x, B) \\ &= \int_B \pi(dx) P^*(x, A) \\ &= \int_S \pi(dx) \mathbb{I}(y \in B) \int_S P^*(x, dy) \mathbb{I}(y \in A) \\ &= \int_S \pi(dx) \psi(x) \int_S P^*(x, dy) \phi(y). \end{aligned}$$

The result holds for indicator functions and can be extended by linearity of the integrals to simple functions $\phi(x) = \sum_k a_k \mathbb{I}(x \in A_k)$ and $\psi(x) = \sum_k b_k \mathbb{I}(x \in B_k)$, where $a_k, b_k \in \mathbb{R}$ and A_k, B_k are Σ -measurable sets. Now let ϕ and ψ be any Σ -measurable functions, then we can decompose them as $\phi = \phi^+ - \phi^-$ and $\psi = \psi^+ - \psi^-$, where $\phi^+, \phi^-, \psi^+, \psi^- \geq 0$ are Σ -measurable functions. Then there exists sequences of non-negative, non-decreasing simple functions $(\phi_n^+), (\phi_n^-), (\psi_n^+)$ and (ψ_n^-) such that $\phi_n^+ \rightarrow \phi^+, \phi_n^- \rightarrow \phi^-, \psi_n^+ \rightarrow \psi^+$ and $\psi_n^- \rightarrow \psi^-$, with convergence interpreted pointwise. The result follows by applying the Monotone Convergence Theorem and linearity of integration. \square

8. APPENDIX C — PROOFS OF EQUILIBRIUM PROPERTIES

In this appendix we prove the results of Section 4. As previously noted, our analysis of the recursion (3.10)-(3.11) assumes the model has a phase structure and the form of the colonisation function essentially excludes the Allee effect. We will work with the more general form of the recursion given by equation (6.30) which reduces to (3.10)-(3.11) when σ has the product form discussed in Section 3. We state our assumptions more fully here together with two other technical assumptions.

- (G) Phase structure: $s(\theta) \geq f(x, \theta)$ for all $(x, \theta) \in [0, \infty) \times \Theta$.
- (H) No Allee effect: Define $\Theta_1 = \{\theta \in \Theta : f(x, \theta) = 0 \text{ for all } (x, \theta) \in [0, \infty)\}$. For all $\theta \in \Theta \setminus \Theta_1$, $f(x, \theta)$ is strictly concave in x .
- (I) $\sup_{\theta \in \Theta} s(\theta) < 1$.

(J) For every $\phi \in C^+(\Omega)$,

$$\sigma(\{(\theta, z) \in \Theta \times \Omega : \phi(z) > 0 \text{ and } P^*(\theta, \Theta_1) < 1 \text{ and } a(\theta) > 0\}) > 0.$$

The importance of Assumptions (G) and (H) has already been discussed in Section 4. The only additional point concerning Assumption (H) to make is that the set Θ_1 is interpreted as the set of patch characteristics which are incompatible with the patch being colonised. Assumption (I) implies that there is no patch characteristic that makes the survival of the local population to the next time period certain. Assumption (J) can be simplified when σ is the product measure $\zeta \times \pi$ since

$$\begin{aligned} & \sigma(\{(\theta, z) \in \Theta \times \Omega : \phi(z) > 0 \text{ and } P^*(\theta, \Theta_1) < 1 \text{ and } a(\theta) > 0\}) \\ &= \left(\int_{\Omega} \mathbb{I}(\phi(z) > 0) \zeta(z) dz \right) \left(\int_{\Theta} \mathbb{I}(P^*(\theta, \Theta_1) < 1 \text{ and } a(\theta) > 0) \pi(d\theta) \right). \end{aligned}$$

As Ω is the support of ζ , it follows $\int_{\Omega} \mathbb{I}(\phi(z) > 0) \zeta(z) dz > 0$ for any $\phi \in C^+(\Omega)$. The conditions on π implied by Assumption (J) are more subtle. If $P^*(\theta, \Theta_1) < 1$ for all $\theta \in \Theta$, then Assumption (J) requires that the set of patch characteristics corresponding to patches with positive area has positive π measure.

Our analysis uses the following notation. For any two functions ϕ and ψ defined on a common domain \mathcal{X} we write $\phi \leq \psi$ if $\phi(x) \leq \psi(x)$ for all $x \in \mathcal{X}$. Similarly, we write $\phi < \psi$ if $\phi \leq \psi$ and $\phi(x) < \psi(x)$ for some $x \in \mathcal{X}$. Finally, we write $\phi \ll \psi$ if $\phi(x) < \psi(x)$ for all $x \in \mathcal{X}$. With slight abuse of notation, we let 0 denote the function that is zero for all $x \in \mathcal{X}$.

Theorem 8.1. *Suppose Assumptions (A)-(D), (G)-(J) hold. For the recursion (6.30) either: (i) 0 is the unique fixed point in $C^+(\Theta \times \Omega)$ or (ii) there are two fixed points in $C^+(\Theta \times \Omega)$ of which one is 0 and the other q^* satisfies $0 \ll q^*$.*

Proof. For any $\phi \in C^+(\Omega)$, defined the recursion

$$q_{t+1}^{\phi}(\theta, z) = \int_{\Theta} s(\eta) q_t^{\phi}(\eta, z) P^*(\theta, d\eta) + \int_{\Theta} f(\phi(z), \eta) (1 - q_t^{\phi}(\eta, z)) P^*(\theta, d\eta). \quad (8.43)$$

Under Assumption (I), the iterations of (8.43) converge to a unique fixed point which we denote by q_{∞}^{ϕ} . Let $(\theta_t^*, t \geq 0)$ be the Markov chain (π, P^*) . We may express q_{∞}^{ϕ} as

$$q_{\infty}^{\phi}(\theta, z) = \sum_{m=1}^{\infty} \mathbb{E} \left\{ f(\phi(z); \theta_m^*) \prod_{n=1}^{m-1} (s(\theta_n^*) - f(\phi(z); \theta_n^*)) \mid \theta_0^* = \theta \right\}. \quad (8.44)$$

Now define the operator $\mathcal{H} : C^+(\Omega) \rightarrow C^+(\Omega)$ by

$$(\mathcal{H}\phi)(z) := \int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) q_\infty^\phi(\tilde{\theta}, \tilde{z}) \sigma(d\tilde{\theta}, d\tilde{z}). \quad (8.45)$$

Note that if ϕ is a fixed point of \mathcal{H} , then q_∞^ϕ is a fixed point of the recursion (6.30). Furthermore, any fixed point of (6.30) can be used to construct a fixed point of \mathcal{H} . The theorem will be proved if we can show that \mathcal{H} satisfies the five conditions of the cone limit set trichotomy [26, Theorem 13].

Lemma 8.2 (Monotonicity). *For any $\phi, \psi \in C^+(\Omega)$, if $\phi \leq \psi$, then $\mathcal{H}\phi \leq \mathcal{H}\psi$.*

Proof. If $q_t^\phi \leq q_t^\psi$, then for all $\eta \in \Theta$

$$\begin{aligned} s(\eta)q_t^\phi(\eta, z) + f(\phi(z), \eta)(1 - q_t^\phi(\eta, z)) &\leq s(\eta)q_t^\phi(\eta, z) + f(\psi(z), \eta)(1 - q_t^\phi(\eta, z)) \\ &\leq s(\eta)q_t^\psi(\eta, z) + f(\psi(z), \eta)(1 - q_t^\psi(\eta, z)), \end{aligned}$$

where the first inequality follows as f is monotone and $\phi \leq \psi$ and the second follows from Assumption (G) and $q_t^\phi \leq q_t^\psi$. After integrating both sides of the inequality over η with respect to $P^*(\theta, d\eta)$, we see $q_{t+1}^\phi \leq q_{t+1}^\psi$. We may take $q_0^\phi = q_0^\psi$ so $q_t^\phi \leq q_t^\psi$ for all $t \geq 0$. Since the iterations of (8.43) converge, $q_\infty^\phi \leq q_\infty^\psi$. Substituting into (8.45), we see that $\mathcal{H}\phi \leq \mathcal{H}\psi$. \square

Lemma 8.3 (Strong sublinearity). *If $\lambda \in (0, 1)$ and $0 \ll \phi$, then $0 \ll \mathcal{H}(\lambda\phi) - \lambda\mathcal{H}\phi$.*

Proof. From the monotonicity property, $q_\infty^{\lambda\phi} \leq q_\infty^\phi$ for all $\lambda \in (0, 1)$. We begin by showing that $\lambda q_\infty^\phi \leq q_\infty^{\lambda\phi}$. As q_∞^ϕ and $q_\infty^{\lambda\phi}$ are fixed points of (8.43),

$$\begin{aligned} q_\infty^{\lambda\phi}(\theta, z) - \lambda q_\infty^\phi(\theta, z) &= \int_{\Theta} s(\eta) [q_\infty^{\lambda\phi}(\eta, z) - \lambda q_\infty^\phi(\eta, z)] P^*(\theta, d\eta) \\ &\quad + \int_{\Theta} (f(\lambda\phi(z), \eta)(1 - q_\infty^{\lambda\phi}(\eta, z)) - \lambda f(\phi(z), \eta)(1 - q_\infty^\phi(\eta, z))) P^*(\theta, d\eta) \\ &= \int_{\Theta} s(\eta) [q_\infty^{\lambda\phi}(\eta, z) - \lambda q_\infty^\phi(\eta, z)] P^*(\theta, d\eta) \\ &\quad + \int_{\Theta} (f(\lambda\phi(z), \eta) - \lambda f(\phi(z), \eta)) (1 - q_\infty^{\lambda\phi}(\eta, z)) P^*(\theta, d\eta) \\ &\quad + \int_{\Theta} \lambda f(\phi(z), \eta) (q_\infty^\phi(\eta, z) - q_\infty^{\lambda\phi}(\eta, z)) P^*(\theta, d\eta). \end{aligned}$$

Since $q_\infty^{\lambda\phi} \leq q_\infty^\phi$ from the monotonicity property,

$$\begin{aligned} q_\infty^{\lambda\phi}(\theta, z) - \lambda q_\infty^\phi(\theta, z) &\geq \int_{\Theta} s(\eta) [q_\infty^{\lambda\phi}(\eta, z) - \lambda q_\infty^\phi(\eta, z)] P^*(\theta, d\eta) \\ &\quad + \int_{\Theta} (f(\lambda\phi(z), \eta) - \lambda f(\phi(z), \eta)) (1 - q_\infty^{\lambda\phi}(\eta, z)) P^*(\theta, d\eta). \end{aligned} \quad (8.46)$$

We can bound q_∞^ϕ from above as follows. As $0 \leq q_t^\phi(\theta, z) \leq 1$ for all $(\theta, z) \in \Theta \times \Omega$, (8.43) implies

$$q_{t+1}^\phi(\theta, z) \leq \int_{\Theta} (s(\eta) \vee f(\phi(z), \eta)) P^*(\theta, d\eta).$$

By Assumption (G), $q_{t+1}^\phi(\theta, z) \leq \sup_{\eta \in \Theta} s(\eta)$ for all $(\theta, z) \in \Theta \times \Omega$. Taking $q_0^\phi = 0$ and noting that the iterations of (8.43) converge to q_∞^ϕ , it follows $q_\infty^\phi(\theta, z) < \sup_{\eta \in \Theta} s(\eta)$ for all $(\theta, z) \in \Theta \times \Omega$. Iterating inequality (8.46) and applying Assumption (I), we see

$$q_\infty^{\lambda\phi}(\theta, z) - \lambda q_\infty^\phi(\theta, z) \geq \left(1 - \sup_{\eta \in \Theta} s(\eta)\right) \int_{\Theta} (f(\lambda\phi(z), \eta) - \lambda f(\phi(z), \eta)) P^*(\theta, d\eta) \quad (8.47)$$

so $\lambda q_\infty^\phi \leq q_\infty^{\lambda\phi}$.

Now if $(\mathcal{H}(\lambda\phi))(z) - \lambda(\mathcal{H}\phi)(z) = 0$ for some $z \in \Omega$, then

$$\int_{\Theta \times \Omega} a(\theta) D(z, \tilde{z}) [q_\infty^{\lambda\phi}(\theta, \tilde{z}) - \lambda q_\infty^\phi(\theta, \tilde{z})] \sigma(d\theta, d\tilde{z}) = 0. \quad (8.48)$$

Equation (8.48) and Assumption (D) together imply that $a \cdot [q_\infty^{\lambda\phi} - \lambda q_\infty^\phi] = 0$, σ -almost everywhere. As $0 \ll \phi$, $f(\lambda\phi(z), \theta) - \lambda f(\phi(z), \theta) > 0$ for all $\theta \in \Theta \setminus \Theta_1$ and all $z \in \Omega$. Using the lower bound (8.47) and Assumption (J), we see that $a \cdot [q_\infty^{\lambda\phi} - \lambda q_\infty^\phi]$ is positive on a set of positive σ -measure. Therefore, equation (8.48) cannot hold for any $z \in \Omega$. \square

Lemma 8.4 (Strong positivity). *If $0 < \phi$, then $0 \ll \mathcal{H}$.*

Proof. If $(\mathcal{H}\phi)(z) = 0$ for some $z \in \Omega$, then

$$\int_{\Theta \times \Omega} a(\theta) D(z, \tilde{z}) q_\infty^\phi(\theta, \tilde{z}) \sigma(d\theta, d\tilde{z}) = 0. \quad (8.49)$$

By Assumption (D), this implies $a(\theta) q_\infty^\phi(\theta, z) = 0$ σ -almost everywhere. Assumption (G) and equation (8.43) imply

$$q_\infty^\phi(\theta, z) \geq \int_{\Theta} f(\phi(z), \eta) P^*(\theta, d\eta). \quad (8.50)$$

Using the lower bound (8.50) and Assumption (J), we see that $a \cdot q_\infty^\phi$ is positive on a set of positive σ -measure. Hence, (8.49) cannot hold for any $z \in \Omega$. \square

Lemma 8.5 (Continuity). *The operator \mathcal{H} is a continuous on $C^+(\Omega)$.*

Proof. For any $\phi, \psi \in C^+(\Omega)$,

$$\sup_{z \in \Omega} |(\mathcal{H}\phi)(z) - (\mathcal{H}\psi)(z)| \leq \bar{D} \int_{\Theta \times \Omega} a(\theta) |q_\infty^\phi(\theta, \tilde{z}) - q_\infty^\psi(\theta, \tilde{z})| \sigma(d\theta, d\tilde{z}) \quad (8.51)$$

by Assumption (D). As q_∞^ϕ and q_∞^ψ are fixed points of (8.43),

$$\begin{aligned} q_\infty^\phi(\theta, z) - q_\infty^\psi(\theta, z) &= \int_{\Theta} (s(\eta) - f(\psi(z), \eta)) (q_\infty^\phi(\theta, z) - q_\infty^\psi(\theta, z)) P^*(\theta, d\eta) \\ &\quad + \int_{\Theta} (f(\phi(z), \eta) - f(\psi(z), \eta)) (1 - q_\infty^\phi(\eta, z)) P^*(\theta, d\eta). \end{aligned}$$

By Assumption (G) and (C),

$$\begin{aligned} \sup_{\theta \in \Theta} |q_\infty^\phi(\theta, z) - q_\infty^\psi(\theta, z)| &\leq \sup_{\theta \in \Theta} \int_{\Theta} s(\eta) |q_\infty^\phi(\theta, z) - q_\infty^\psi(\theta, z)| P^*(\theta, d\eta) + L |\phi(z) - \psi(z)| \\ &\leq \left(\sup_{\theta \in \Theta} s(\theta) \right) \left(\sup_{\theta \in \Theta} |q_\infty^\phi(\theta, z) - q_\infty^\psi(\theta, z)| \right) + L |\phi(z) - \psi(z)|. \end{aligned}$$

By Assumption (I),

$$\sup_{\theta \in \Theta} |q_\infty^\phi(\theta, z) - q_\infty^\psi(\theta, z)| \leq L \left(1 - \sup_{\theta \in \Theta} s(\theta) \right) |\phi(z) - \psi(z)|. \quad (8.52)$$

Substituting this into inequality (8.51), we see that \mathcal{H} is a (Lipschitz) continuous operator on $C^+(\Omega)$. \square

Lemma 8.6 (Order compactness). *For any $\psi_1, \psi_2 \in C^+(\Omega)$, \mathcal{H} maps the set $\{\phi \in C^+(\Omega) : \psi_1 \leq \phi \leq \psi_2\}$ to a relatively compact set.*

Proof. As $q_\infty^\phi(\theta, z) \leq 1$ for all $(\theta, z) \in \Theta \times \Omega$ and $D(\cdot, \cdot)$ is uniformly bounded and equicontinuous by Assumption (D), the image of $\{\phi \in C^+(\Omega) : \psi_1 \leq \phi \leq \psi_2\}$ under \mathcal{H} is a set of uniformly continuous functions. Order compactness now follows from the Arzelà-Ascoli theorem. \square

We may now apply the cone limit set trichotomy. We have seen in the proof of strong sublinearity that for any $\phi \in C^+(\Omega)$, $q_\infty^\phi(\theta, z) < \sup_{\eta \in \Theta} s(\eta)$ for all $(\theta, z) \in \Theta \times \Omega$. Substituting this bound into equation (8.45) shows that \mathcal{H} is a bounded operator. This excludes the possibility of an orbit of \mathcal{H} being unbounded. Therefore, either (i) each orbit of \mathcal{H} converges to 0, the unique fixed point of \mathcal{H} , or (ii) each nonzero orbit converges to $q^* \gg 0$; the unique nonzero fixed point of \mathcal{H} . This completes the proof of Theorem 8.1. \square

To determine which of the two possibilities from Theorem 8.1 occurs, we need to define the operator $\mathcal{A} : C^+(\Omega) \rightarrow C^+(\Omega)$,

$$(\mathcal{A}\phi)(z) = \int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \phi(\tilde{z}) \sum_{m=1}^{\infty} \mathbb{E} \left\{ f'(0; \theta_m^*) \prod_{n=1}^{m-1} s(\theta_n^*) \mid \theta_0^* = \tilde{\theta} \right\} \sigma(d\tilde{z}, d\tilde{\theta}).$$

The spectral radius of \mathcal{A} is denoted $r(\mathcal{A})$.

Theorem 8.7. *Suppose Assumptions (A)-(D), (G)-(J) hold. If $r(\mathcal{A}) \leq 1$, then 0 is the unique fixed point of the recursion (6.30). If $r(\mathcal{A}) > 1$, then recursion (6.30) has a non-zero fixed point.*

Proof. By Assumptions (C), (G) and (H)

$$\phi(z) \mathbb{E} \left\{ f'(0; \theta_m^*) \prod_{n=1}^{m-1} s(\theta_n^*) \mid \theta_0^* = \theta \right\} - q_{\infty}^{\phi}(\theta, z) \geq \mathbb{E} \{ \phi(z) f'(0; \theta_1^*) - f(\phi(z); \theta_1^*) \mid \theta_0^* = \theta \}$$

so

$$(\mathcal{A}\phi)(z) - (\mathcal{H}\phi)(z) \geq \int_{\Theta \times \Omega} a(\tilde{\theta}) D(z, \tilde{z}) \mathbb{E} \left\{ \phi(z) f'(0; \theta_1^*) - f(\phi(z); \theta_1^*) \mid \theta_0^* = \tilde{\theta} \right\} \sigma(d\tilde{z}, d\tilde{\theta}).$$

Now $\phi(z) f'(0; \theta) - f(\phi(z); \theta) > 0$ if $\phi(z) > 0$ and $\theta \in \Theta \setminus \Theta_1$. By Assumption (D), $D(z, \tilde{z}) > 0$ for all $(z, \tilde{z}) \in \Omega \times \Omega$. Assumption (J) now implies that if $\phi \neq 0$, then

$$(\mathcal{A}\phi)(z) - (\mathcal{H}\phi)(z) > 0,$$

for all $z \in \Omega$. If $\phi = 0$, then by substitution $\mathcal{A}\phi = \mathcal{H}\phi = 0$. By [41, Lemma A.1], it follows that if $r(\mathcal{A}) \leq 1$, then 0 is the unique fixed point of \mathcal{H} and hence is the unique fixed point of the recursion (3.10)-(3.11).

As in [41, Lemma A.2], if $r(\mathcal{A}) > 1$, then there exists a $\psi \in C^+(\Omega)$ such that $\psi \leq \mathcal{H}\psi$. This implies that the set $\{\phi \in C^+(\Omega) : \psi \leq \phi\}$ is invariant under \mathcal{H} . The Schauder fixed point theorem [27, Theorem 5.1.2] shows that \mathcal{H} has a fixed point in $\{\phi \in C^+(\Omega) : \psi \leq \phi\}$. This non-zero fixed point must be unique by Theorem 8.1. \square

Theorem 8.8. *Suppose Assumptions (A)-(D), (G)-(J) hold. If $r(\mathcal{A}) \leq 1$, then all trajectories of the recursion (6.30) converge to 0. If $r(\mathcal{A}) > 1$ and $\frac{\partial \mu_0}{\partial \sigma}(\theta, z) > \epsilon$ for all $(\theta, z) \in \Theta \times \Omega$ and some $\epsilon > 0$, then trajectory of the recursion (6.30) converge to the non-zero fixed point.*

Proof. Under Assumptions (C) and (G), we can use a similar argument to that used in the proof of Lemma 8.2 to show that recursion (6.30) has the monotonicity property; if $\frac{\partial \tilde{\mu}_0}{\partial \sigma}$ and $\frac{\partial \mu_0}{\partial \sigma}$ are two initial conditions such that $\frac{\partial \tilde{\mu}_0}{\partial \sigma} \leq \frac{\partial \mu_0}{\partial \sigma}$ in the partial ordering on $C^+(\Omega)$, then under recursion (6.30) $\frac{\partial \tilde{\mu}_t}{\partial \sigma} \leq \frac{\partial \mu_t}{\partial \sigma}$ for all $t \geq 0$. The proof then follows the arguments of [10, Theorems 5.1 and 5.3] as used in [39, Theorem 3]. \square

Proof of Theorem 4.1. The result follows from Theorems 8.1, 8.7 and 8.8 with σ replaced by the product measure $\zeta \times \pi$. \square

Proof of Theorem 4.2. Assuming $f(x; \theta) = s(\theta)\bar{f}(x)$, equation (8.44) becomes

$$q_\infty^\phi(\theta, z) = \sum_{m=1}^{\infty} \bar{f}(\phi(z))(1 - \bar{f}(\phi(z)))^{m-1} \mathbb{E} \left\{ \prod_{n=1}^m s(\theta_n^*) \mid \theta_0^* = \theta \right\}. \quad (8.53)$$

Integrating $a(\theta)q_\infty^\phi(\theta, z)$ with respect to π shows that

$$\int_{\Theta} a(\theta)q_\infty^\phi(\theta, z)\pi(d\theta) = \sum_{m=1}^{\infty} \bar{f}(\phi(z))(1 - \bar{f}(\phi(z)))^{m-1} \mathbb{E} \left\{ a(\theta_{m+1}) \prod_{n=1}^m s(\theta_n) \right\}. \quad (8.54)$$

Substituting equation (8.54) into equation (8.45) and using the product form for σ yields

$$(\mathcal{H}\phi)(z) = \int_{\Omega} D(z, \tilde{z}) \sum_{m=1}^{\infty} \bar{f}(\phi(z))(1 - \bar{f}(\phi(z)))^{m-1} \mathbb{E} \left\{ a(\theta_{m+1}) \prod_{n=1}^m s(\theta_n) \right\} \zeta(\tilde{z}) d\tilde{z}. \quad (8.55)$$

As \mathcal{H} depends on (π, P) only through the sequence $\mathbb{E}(a(\theta_{m+1}) \prod_{n=1}^m s(\theta_n))$, $m \geq 1$ so must its fixed point. \square

Proof of Theorem 4.3. Let $\tilde{\mathcal{H}}$ be the operator obtained by replacing θ_n with $\tilde{\theta}_n$ in equation (8.55). From (4.17), it follows that for any $\phi \in C^+(\Omega)$

$$\tilde{\mathcal{H}}\phi \leq \mathcal{H}\phi.$$

If $\tilde{\phi}^*$ is a fixed point of $\tilde{\mathcal{H}}$, then $\tilde{\phi}^* = \tilde{\mathcal{H}}\tilde{\phi}^* \leq \mathcal{H}\tilde{\phi}^*$. Therefore, the set $\{\phi \in C^+(\Omega) : \tilde{\phi}^* \leq \phi\}$ is invariant under \mathcal{H} and by the Schauder fixed point theorem \mathcal{H} has a fixed point in this set. As \mathcal{H} has a unique non-zero fixed point (Theorem 8.1), $\tilde{\phi}^* \leq \phi^*$. Now by equation (8.54)

$$\begin{aligned} \int_{\Theta} a(\theta)\tilde{q}^*(\theta, z)\tilde{\pi}(d\theta) &= \sum_{m=1}^{\infty} \bar{f}(\phi(z))(1 - \bar{f}(\tilde{\phi}^*(z)))^{m-1} \mathbb{E} \left\{ a(\tilde{\theta}_{m+1}) \prod_{n=1}^m s(\tilde{\theta}_n) \right\} \\ &\leq \sum_{m=1}^{\infty} \bar{f}(\phi(z))(1 - \bar{f}(\tilde{\phi}^*(z)))^{m-1} \mathbb{E} \left\{ a(\theta_{m+1}) \prod_{n=1}^m s(\theta_n) \right\} \end{aligned}$$

from inequality (4.17). Now let $\alpha_m = \mathbb{E}(a(\theta_{m+1}) \prod_{n=1}^m s(\theta_n)) - \mathbb{E}(a(\theta_{m+2}) \prod_{n=1}^{m+1} s(\theta_n))$. Then

$$\begin{aligned} \int_{\Theta} a(\theta) \tilde{q}^*(\theta, z) \tilde{\pi}(d\theta) &\leq \sum_{m=1}^{\infty} \bar{f}(\tilde{\phi}^*(z)) (1 - \bar{f}(\tilde{\phi}^*(z)))^{m-1} \left\{ \sum_{r=m}^{\infty} \alpha_r \right\} \\ &= \sum_{r=1}^{\infty} (1 - (1 - \bar{f}(\tilde{\phi}^*(z)))^r) \alpha_r \end{aligned}$$

As $\tilde{\phi}^* \leq \phi^*$ and \bar{f} is increasing, $(1 - (1 - \bar{f}(\tilde{\phi}^*(z)))^r) \leq (1 - (1 - \bar{f}(\phi^*(z)))^r)$ for all $r \geq 1$ and all $z \in \Omega$. Therefore,

$$\begin{aligned} \int_{\Theta} a(\theta) \tilde{q}^*(\theta, z) \tilde{\pi}(d\theta) &\leq \sum_{r=1}^{\infty} (1 - (1 - \bar{f}(\phi^*(z)))^r) \alpha_r \\ &= \sum_{m=1}^{\infty} \bar{f}(\phi(z)) (1 - \bar{f}(\phi^*(z)))^{m-1} \mathbb{E} \left\{ a(\theta_{m+1}) \prod_{n=1}^m s(\theta_n) \right\} \\ &= \int_{\Theta} a(\theta) q^*(\theta, z) \pi(d\theta). \end{aligned}$$

□

Proof of Lemma 4.4. From the general form of Hölder's inequality

$$\mathbb{E} \left(\prod_{n=0}^m s(\theta_n) \right) \leq \prod_{n=0}^m \mathbb{E} (s(\theta_n)^{m+1})^{1/(m+1)}.$$

As $(\theta_t, t \geq 0)$ is assumed stationary $\mathbb{E}(s(\theta_t)^{m+1}) = \mathbb{E}(s(\theta_0)^{m+1})$ for $t = 1, \dots, m$. Therefore,

$$\mathbb{E} \left(\prod_{n=0}^m s(\theta_n) \right) \leq \mathbb{E}(s(\theta_0)^{m+1}).$$

□

8.1. Numerical approximation of the fixed point. To compute $\int_{\Theta} q^*(\theta, z) \pi(d\theta)$, we first numerically determined the non-zero fixed point the operator \mathcal{H} from (8.45) employing the product form of σ and the expression for $\int_{\Theta} q^{\phi}(\theta, z) \pi(d\theta)$ given by (8.54). This was done by fixed point iteration of an approximation to the operator \mathcal{H} where we (i) approximated the integral with respect to ζ by a Reimann sum with 500 terms, (ii) truncated the infinite sum in (8.54) to 1000 terms, and (iii) approximated the moments $\mathbb{E}(\prod_{n=1}^m s(\theta_n))$ by simulating 1000 sample paths of the survival process. The fixed point of \mathcal{H} was then substituted into equation (8.54) to give an approximation to the limiting probability of the patch being occupied.

REFERENCES

- [1] Akçakaya HR, and Ginzburg LR (1991) Ecological risk analysis for single and multiple populations, pages 78-87 in Species Conservation: A Population Biological Approach (Seitz A and Loescheke V, eds.), Birkhauser, Basel
- [2] Akçakaya HR, Radeloff VC, Mladenoff DJ and He HS (2004) Integrating landscape and metapopulation modeling approaches: Viability of the sharp-tailed grouse in a dynamic landscape, Conservation Biology, 18, 526-537
- [3] Bassch A, Tischew S and Bruelheide H (2010) Twelve years of succession on sandy substrates in a post-mining landscape: a Markov chain analysis, Ecological Applications, 20, 1136-1147
- [4] Barker WL (1989) A review of models of landscape change, Landscape Ecology, 2, 111-133
- [5] Balzter H (2000) Markov chain models for vegetation dynamics, Ecological Modelling, 126, 139-154.
- [6] Barbour AD, McVinish R and Pollett PK (2015) Connecting deterministic and stochastic metapopulation models, Journal of Mathematical Biology, 71, 1481-1504
- [7] Bebbington M, Pollett PK and Zheng X (1995) Dual constructions for pure-jump Markov processes, Markov Processes and Related Fields, 1, 513-558
- [8] Boyle OD, Menges ES and Waller DM (2003) Dances with fire: Tracking metapopulation dynamics of *Polygonella Basiramia* in Florida scrub (USA), Folia Geobotanica, 38, 255-262
- [9] Brachet S, Oliveria I, Godelle B, Klein E, Frascaria-Lacoste N and Gouyon P-H (1999) Dispersal and Metapopulation Viability in a Heterogeneous Landscape, Journal of Theoretical Biology, 198, 479-495
- [10] Busenberg SN, Iannelli M and Thieme HR (1991) Global behavior of an age-structured epidemic model, SIAM Journal on Mathematical Analysis, 22, 1065-1080
- [11] Chesson P (1984) Persistence of a Markovian population in a patchy environment, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 66, 97-107
- [12] Courchamp, F., Berec, L. and Gascoigne, J. (2008) Allee effects in ecology and conservation, Oxford University Press, Oxford

- [13] Day JR and Possingham HP (1995) A stochastic metapopulation model with variability in patch size and position, *Theoretical Population Biology*, 48, 333-360
- [14] DeWoody YD, Feng Z and Swihart RK (2005) Merging spatial and temporal structure within a metapopulation model, *American Naturalist*, 166, 42-55
- [15] Diekmann O, Heesterbeek JAP, and Roberts MG (2010) The construction of next-generation matrices for compartmental epidemic models, *Journal of the Royal Society Interface*, 7, 873-885
- [16] Dolrenry S, Stenglein J, Hazzah L, Lutz RS and Frank L (2014) A metapopulation approach to African lion (*Panthera leo*) conservation, *PLoS ONE*, 9, e88081
- [17] Durrett R and Levin SA (1994) The importance of being discrete (and spatial), *Theoretical Population Biology*, 46, 363-394.
- [18] Franc A (2004) Metapopulation dynamics as a contact process on a graph, *Ecological Complexity*, 1, 49-63
- [19] George DB, Webb CT, Pepin KM, Savage LT and Antolin MF (2013) Persistence of black-tailed prairie-dog populations affected by plague in northern Colorado, USA, *Ecology*, 94, 1572-1583.
- [20] Gregg L, and Niemuth ND (2000) The history, status, and future of the sharp-tailed grouse in Wisconsin, *The Passenger Pigeon* 62:159-174
- [21] Hanski I (1991) Single-species metapopulation dynamics: concepts, models and observations, *Biological Journal of the Linnean Society*, 42, 17-38.
- [22] Hanski I (1994) A practical model of metapopulation dynamics, *Journal of Animal Ecology*, 63, 151-162
- [23] Hanski I (1999) Habitat connectivity, habitat continuity, and metapopulations in dynamic landscapes, *Oikos*, 87, 209-219
- [24] Hanski I and Ovaskainen O (2003) Metapopulation theory for fragmented landscapes, *Theoretical Population Biology*, 64, 119-127
- [25] Hill MF and Caswell H (2001) The effects of habitat destruction in finite landscapes: A chain-binomial metapopulation model, *Oikos*, 93, 321-331
- [26] Hirsch MW and Smith H (2005) Monotone maps: A review, *Journal of Difference Equations and Applications*, 11, 379-398
- [27] Istratescu VI (1981) *Fixed Point Theory: An Introduction*, D. Reidel, Dordrecht.

- [28] Johansson V, Ranius T and Snäll T (2012) Epiphyte metapopulation dynamics are explained by species traits, connectivity, and patch dynamics, *Ecology*, 93, 235-241
- [29] Kallenberg O (2002) Foundations of modern probability, 2nd edn. Springer, New York
- [30] Kelly FP (1979) Reversibility and Stochastic Networks, Wiley, Chichester
- [31] Keymer JE, Marquet PA, Velasco-Hernandez JX and Levin SA (2000) Extinction thresholds and metapopulation persistence in dynamic landscapes, *The American Naturalist*, 156, 478-494
- [32] Léonard C (1990) Some epidemic systems are long range interacting particle systems, pages 170-183 in *Stochastic Processes in Epidemic Theory*. (Gabriel J-P, Lefèvre C and Picard P, eds.), Springer
- [33] Levins R (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control, *Bulletin of the Entomological Society of America*, 15: 237-240
- [34] Liggett TM (2005) *Interacting Particle Systems*, Springer, Germany.
- [35] Logofet DO and Lesnaya EV (2000) The mathematics of Markov models: what Markov chains can really predict in forest successions, *Ecological Modelling*, 186, 285-298
- [36] MacPherson JL and Bright PW (2011) Metapopulation dynamics and a landscape approach to conservation of lowland water voles (*Arvicola amphibius*), *Landscape Ecology*, 26, 1395-1404
- [37] McKinlay S and Borovkov K (2016) On explicit form of the stationary distributions for a class of bounded Markov chains, *Journal of Applied Probability*, 53, 231-243.
- [38] McVinish R and Pollett PK (2012) The limiting behaviour of a mainland-island metapopulation, *Journal of Mathematical Biology*, 64, 775-801
- [39] McVinish R and Pollett PK (2013) The limiting behaviour of a stochastic patch occupancy model, *Journal of Mathematical Biology*, 67, 693-716
- [40] McVinish R and Pollett PK (2013) The deterministic limit of a stochastic logistic model with individual variation, *Mathematical Biosciences*, 241, 109-114
- [41] McVinish R and Pollett PK (2014) The limiting behaviour of Hanski's incidence function metapopulation model, *Journal of Applied Probability*, 51, 297-316

- [42] Metz JAJ and Gyllenberg M (2001) How Should We Define Fitness in Structured Metapopulation Models? Including an Application to the Calculation of Evolutionarily Stable Dispersal Strategies, *Proceedings of the Royal Society of London Series B Biological Sciences*, 268, 499-508
- [43] Meyn SP and Tweedie RL (1996) *Markov Chains and Stochastic Stability*, Springer-Verlag, London.
- [44] Miller AD and Chesson P (2009) Coexistence in disturbance-prone communities: How a resistance-resilience trade-off generates coexistence via the storage effect, *American Naturalist*, 173, E30-E43
- [45] Moilanen A and Nieminen M (2002) Simple connectivity measures in patial ecology, *Ecology*, 83, 1131-1145.
- [46] Neal P (2014) Endemic behaviour of SIS epidemics with general infectious period distributions, *Advances in Applied Probability*, 46, 241–255
- [47] O’Cinneide CA (1990) Characterization of phase-type distributions, *Communications in Statistics: Stochastic Models*, 6, 1-57
- [48] Ovaskainen O (2001) The quasistationary distribution of the stochastic logisitic model, *Journal of Applied Probability*, 38, 898-907
- [49] Ovaskainen O and Cornell SJ (2006) Asymptotically exact analysis of stochastic metapopulation dynamics with explicit spatial structure, *Theoretical Population Biology*, 69, 13-33
- [50] Ovaskainen O and Hanski I (2001) Spatially structured metapopulation models: global and local assessment of metapopulation capacity, *Theoretical Population Biology*, 60, 281-302.
- [51] Pulsford SA, Lindenmayer DB and Driscoll DA (2016) A succession of theories: purging redundancy from disturbance theory, *Biological Reviews*, 91, 148-167.
- [52] Ranga Rao, R (1962) Relations between weak and uniform convergence of measures with applications, *Annals of Mathematical Statistics*, 33, 659-680
- [53] Ranius T, Bohman P, Hedgren O, Wikars L-O and Caruso A (2014) Metapopulation dynamics of a beetle species confined to burned forest sites in a managed forest region, *Ecography* 37, 797-804
- [54] Reigada C, Schreiber SJ, Altermatt F and Holyoak M (2015) Metapopulation dynamics on ephemeral patches, *American Naturalist*, 185, 183-195.

- [55] Ross JV (2006) A stochastic metapopulation model accounting for habitat dynamics, *Journal of Mathematical Biology*, 52, 788-806
- [56] Roxburgh SH, Shea K and Wilson JB (2004) The intermediate disturbance hypothesis: Patch dynamics and mechanisms of species coexistence, *Ecology*, 85, 359-371
- [57] Shaked M, Shanthikumar JG (2007) *Stochastic Orders*, Springer, New York.
- [58] Shea K and Chesson P (2002) Community ecology theory as a framework for biological invasions, *Trends in Ecology & Evolution*, 17, 170-176
- [59] Turner MG, Baker WL, Peterson C, and Peet RK (1998) Factors influencing succession: lessons from large, infrequent natural disturbances, *Ecosystems*, 1, 511-523
- [60] Usher MB (1979) Markovian approaches to ecological succession, *Journal of Animal Ecology*, 48, 413-426.
- [61] van den Driessche P and Watmough J (2002) Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, *Mathematical Biosciences*, 180, 29-48
- [62] van Teeffelen AJA, Vos CC and Opdam P (2012) Species in a dynamic world: Consequences of habitat dynamics on conservation planning, *Biological Conservation*, 153, 239-253
- [63] Verheyen K, Vellend M, Van Calster H, Peterken G and Hermy M (2004) Metapopulation dynamics in changing landscapes: A new spatially realistic model for forest plants, *Ecology*, 85, 3302-3312
- [64] Weiss GH and Dishon M (1971) On the asymptotic behaviour of the stochastic and deterministic models of an epidemic, *Mathematical Biosciences*, 11, 261-265
- [65] Wilcox C, Cairns BJ and Possingham HP (2006) The role of habitat disturbance and recovery in metapopulation persistence, *Ecology*, 87, 855-863
- [66] Xu D, Feng Z, Allen LJS, Shiwart RK (2006) A spatially structured metapopulation model with patch dynamics, *Journal of Theoretical Biology*, 239, 469-481